2006-01-01

Potential of Vibrational Spectroscopy in the Diagnosis of Human Tumours.

Eoghan O'Faolain
Technological University Dublin

Recommended Citation

O'Faolain, Eoin. (2006). *Potential of vibrational spectroscopy in the diagnosis of human tumours*.Doctoral thesis. Technological University Dublin. doi:10.21427/D7DW31

School of Physics
Dublin Institute of Technology
Kevin Street, Dublin 8

Potential of Vibrational Spectroscopy in the Diagnosis of Human Tumours

Eoghan Ó Faoláin B.Sc. (Hons.)

May 2006

Thesis submitted for PhD examination
Dublin Institute of Technology

Supervised by

Dr. Fiona M. Lyng (DIT)
Dr. Hugh J. Byrne (DIT)

ABSTRACT

Just fewer than 20,000 people are annually diagnosed with some form of cancer in Ireland, and one in three people are likely to contract some form of cancer by age 74. With the number of cases increasing at an annual rate of 2%, the early detection and treatment of cancer is becoming increasingly important. Both IR and Raman spectroscopy offer the potential for real time, quantitative detection of cancer and even precancer. This study investigates the potential of Raman and Fourier transform infrared, both benchtop and synchrotron spectroscopies for the detection of cervical cancer. The tissue was classified and its various spectral contributions were associated with the chemical bonds in the tissue's constituent components. Bands have been attributed to biological compounds such as carbohydrates, proteins, lipids and nucleic acids. Formalin fixed paraffin preserved (FFPP) tissue sections were found to be sufficiently similar to unprocessed tissue to be used in the study. The study identified problems with residual wax in these FFPP sections and recommended the use of hexane as a dewaxing solvent. Spectra recorded from normal cervical tissue were compared to spectra from cervical intraepithelial neoplasia (CIN) and invasive carcinoma. Significant spectral differences were observed between normal and cancerous, as well as neoplastic tissue. Principal component analysis and linear discriminant analysis yielded sensitivities of 99%, 99%, and 98.5% and specificities of 100%, 99% and 99% for normal epithelial tissue, CIN, and invasive carcinoma respectively. These high values of sensitivity and specificity achieved confirms vibrational spectroscopy as a useful tool in the detection of cancer, as a well as useful tool to understand the biochemical changes in tissue with the onset of disease.

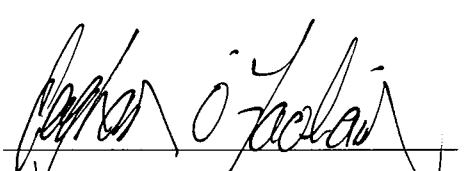
DECLARATION

I certify that this thesis which I now submit for examination for the award of Ph.D., is my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for an award in any other Institute or University.

The work reported on in this thesis conforms to the principles and requirements of the Institute's guidelines for Ethics in Research.

The Institute has permission to keep, to lend or to copy this thesis in whole or in part on condition that any such use of the material of the thesis be duly acknowledged.

Signature  Date Oct '06

Eoghan Ó Faoláin B.Sc. (Hons.)

ACKNOWLDGEMENTS

Thanks to all the technical staff at the Focas Institute over the years, Laura Evans, Garrett Farrell, Anne Shanahan and Andrew Hartnett. I've lost count of how often you got me out of trouble. And of course a special thanks to Louisa Hartnett for her help with all my random queries and requests.

Also my sincerest thanks to Dr. Helen Lambkin (Dept. of Biological Sciences, D.I.T. Kevin St.) for her help with the immunohistochemical study and constant insight into all things biological, to whom this physicist is a constant source of bewilderment.

A big thank you to Dr. Mark Tobin, Dr. Marisa Martin-Fernandez, and Fariba Bahrami for the warm welcome in the U.K. Thanks to Mark and Fariba for all their help on the beamline and subsequently!

Without co-operation of the staff in the Histology Department in the NMH Holles St. this project would not be possible, in particular Mary Hunter, Joe Byrne and Dr. Peter Kelehan. Dr. Kelehan, thank you for having the inquisitive passion for research and long may you continue to encourage it. Mary and Joe, you two are fantastic, thank you for your knowledge, patience and generosity with your time.

Thanks also to Jennifer Conroy, for our preliminary investigations we carried out on component molecules for her undergraduate project.

Also a huge thank you to Dr. Brendan Duffy for all his help with the multivariate analysis of the thousands of spectra generated during this PhD. I think we both know I couldn't have done it without you.

A huge thank you must go to Dr. Hugh Byrne and Dr. Mary McNamara for their selfless generosity both to myself and all the research students in the Focas Institute.

Dr. Fiona Lyng, what can I say? I wouldn't be here without you. I shudder when I look back over this *long* journey. The ups (and there were many) and downs (maybe a few), trials (almost) and tribulations, I don't know how you put up with me. I can't even begin to thank you.

And finally, I must thank my parents. Mam and Dad, always supportive, encouraging and optimistic. Thank you for 28 years of minding me. I promise I will get a job one day, it all started with that freezer!

Dá fhada an lá tagann an tráthnóna!

If I have forgotten anyone my apologies, but thanks to everyone who helped me during this PhD.

ABBREVIATIONS

A	Adenine
ALA	Alanine
ARG	Arginine
ASCUS	Atypical Squamous Cells of Undetermined Significance
ATR	Attenuated Total Reflectance
C	Cytosine
CA	Cervical Adenocarcinoma
CIN	Cervical Intraepithelial Tissue
DNA	Deoxyribonucleic Acid
FFPP	Formalin Fixed Paraffin Preserved
FTIR	Fourier Transform Infrared
G	Guanine
GLY	Glycine
H+E	Haematoxylin and Eosin
HMAR	Heat Mediated Antigen Retrieval
HSIL	High-Grade Squamous Intraepithelial Lesion
LDA	Linear Discriminant Analysis
LEU	Leucine
LSIL	Low-Grade Squamous Intraepithelial Lesion
LYS	Lysine
NADH	Nicotinamide Adenine Dinucleotide
NIR	Near-Infrared

NMH	National Maternity Hospital
OCT	Optimal Cutting Temperature
PBS	Phosphate Buffered Saline
PCA	Principal Components Analysis
RNA	Ribonucleic Acid
SEN	Sensitivity
SER	Serine
SIL	Squamous Intraepithelial Lesion
SPEC	Specificity
SR-FTIR	Synchrotron Radiation for Fourier Transform Infrared
SRS	Synchrotron Radiation Source
T	Tyrosine
PCs	Principal Components

List of Figures

List of Tables

Chapter 1

General Introduction

1.1 Introduction to Cancer

At the mention of the word spectroscopy, the scientific community tends to think of investigations into the chemical structure of polymers, nanotubes and other novel molecules. However vibrational spectroscopy can also be employed to investigate the chemical composition of cells and tissues. Examined simplistically, the human body is composed primarily of water, proteins, nucleic acids, lipids and carbohydrates. Some of these in turn are composed of amino acids, purines and pyrimidines. It is these molecules that bond together and form the complex structure of higher order molecules such as proteins and nucleic acids.

All of these molecules can be probed and identified using spectroscopic techniques. Their presence, quantities and even their structure can be observed. It is not unreasonable to believe that any changes in the body leading to disease (such as cancer) arise due to some biochemical changes in one or all of these components. This rapidly evolving interface between physical sciences and the area broadly termed "biochemical science", offers exciting opportunities to better understand and detect disease early on a cellular level.

1.1.1 Cell Biology

Cells grow and divide to produce daughter cells, which are likewise capable of replicating themselves. Cells typically have diameters of the order of 10μm and were first seen over 300 years ago shortly after the construction of the first microscope. In all eukaryotic

(nucleus containing) cells, the inner cellular mass is partitioned into a membrane bound, spherical body called the nucleus and an outer surrounding cytoplasm which contains the rest of the organelles such as the endoplasmic reticulum and the mitochondria (Fig.1.1). Cellular DNA is located in the nucleus in the form of rods known as chromosomes. Cells are chemically very sophisticated – even the simplest cells contain about 1000 different molecules. These molecules include various sugars, amino acids, fatty acids, proteins and nucleic acids. There are over 200 different types of cells in the human body, such as heart cells, muscle cells, liver cells, retina cells, red and white blood cells.

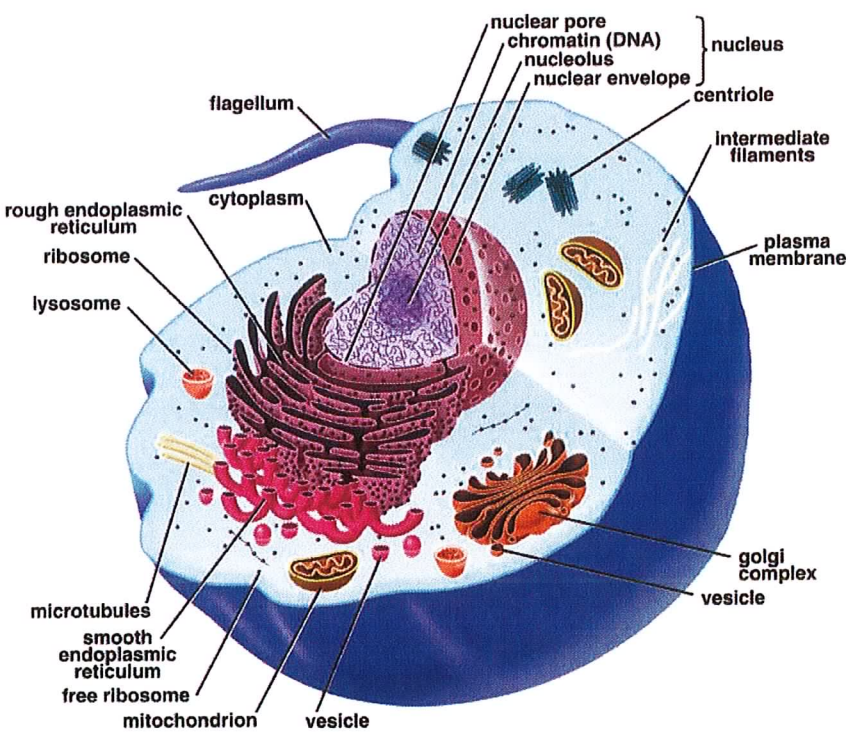


Figure 1.1 Schematic of human cell (Caron 2004)

These cells are assembled into four types of tissue: epithelial, connective, muscular and nerve (Fig.1.2). Most tissues contain a mixture of cell types.

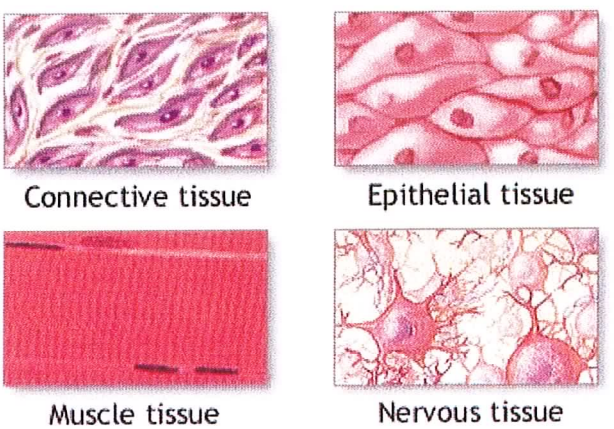


Figure 1.2 Four tissue types found in the human body

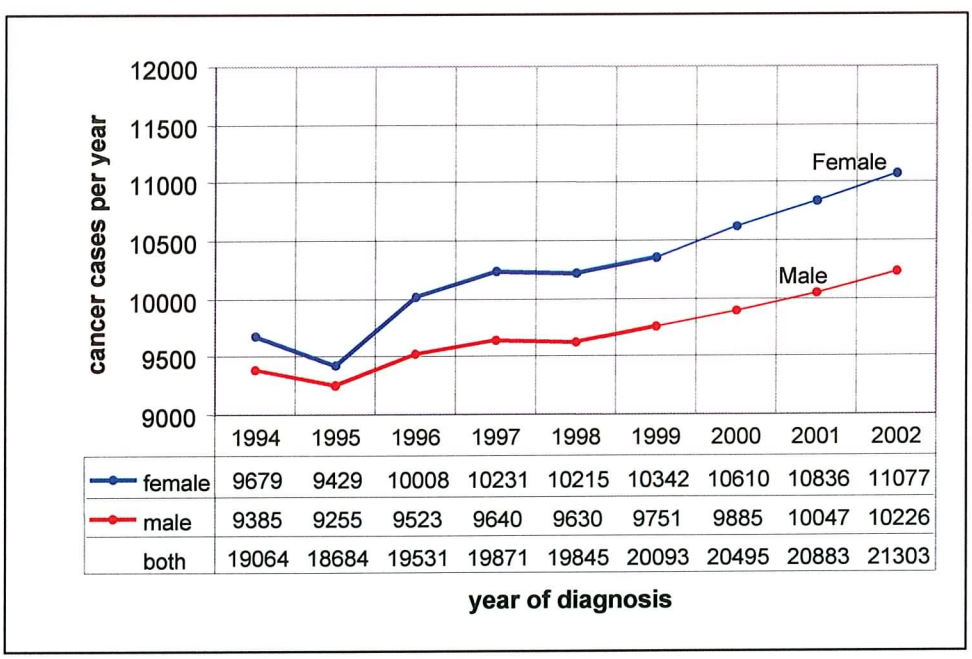
(American Accreditation HealthCare Commission 2004)

Epithelial tissues are composed of closely aggregated polyhedral cells with very little intercellular substance. These epithelial cells form coherent cell sheets that line the inner and outer surfaces of the body. Connective tissue is characterized by the abundance of intercellular material produced by its cells. This material is composed mostly of a network of tough protein fibers made up of collagen and elastin, embedded in a polysaccharide gel. Connective tissue (found in the sub-epithelial layer of many organ systems) contains several types of cells, each having its own morphological and functional characteristics. The most common cell found in connective tissue is the fibroblast. Nerve tissue is found all over the body as a communication network. Nerve tissue consists of nerve cells or neurons and several types of glial cells or neuroglia, which support neurons. Muscular tissues can be divided into three main types, skeletal, smooth and cardiac, each composed of cells with

different function and appearance. The four tissues, in association with one another and in variable proportions, form different organs and systems of the body.

1.1.2 Cancer

Approximately one-quarter of all deaths in Ireland are due to cancer, which makes it the single most important cause of death. The total number of cancer cases is estimated to have increased at an annual rate of 1.8% for women and 1.1% for men between 1994 and 2001. This is an increase of about 300 cases each year. Between 1994 and 2001, an average of 19,808 cancer cases was registered each year (National Cancer Registry 2005), (Fig. 1.3).



	1994	1995	1996	1997	1998	1999	2000	2001	2002
female	9679	9429	10008	10231	10215	10342	10610	10836	11077
male	9385	9255	9523	9640	9630	9751	9885	10047	10226
both	19064	18684	19531	19871	19845	20093	20495	20883	21303

Figure 1.3 Irish cancer statistics (The National Cancer Registry)

However, cancer is not a new disease. Hippocrates is reported to have distinguished benign from malignant growths in 460 AD (Gowri 2000). He introduced the term "karkinos", from which the word "carcinoma" is derived. Cancer is a complex family of diseases and carcinogenesis – the turning of a normal cell into a cancer cell – is a complex multistep process.

Figure 1.4 shows a simplified diagram of cancer progression. In a simple description, a tumour (or neoplasm, which means "new growth") begins when some cell within a normal population sustains a mutation that increases its propensity to proliferate when it would normally rest as shown in Figure 1.4(a). The altered cell and its daughters look normal but they grow and divide too much. This condition is termed *hyperplasia* (Figure 1.4(b)). After some time, which could be years, one of these cells suffers another mutation that further loosens control on cell growth. This cell proliferates rapidly and the offspring of this cell appears abnormal in shape and orientation, as shown in Figure 1.4(c). This state is known as *dysplasia*. Again, after some time, a mutation occurs that alters cell behaviour. The affected cells become even more abnormal in growth and appearance. If the tumour is contained within original tissue, it is called *in situ* cancer (Figure 1.4(d)). This tumour may remain contained indefinitely but often the cells acquire additional mutations, continue to grow and divide and invade underlying tissue. At this point the mass is considered to have become malignant (Figure 1.4(e)).

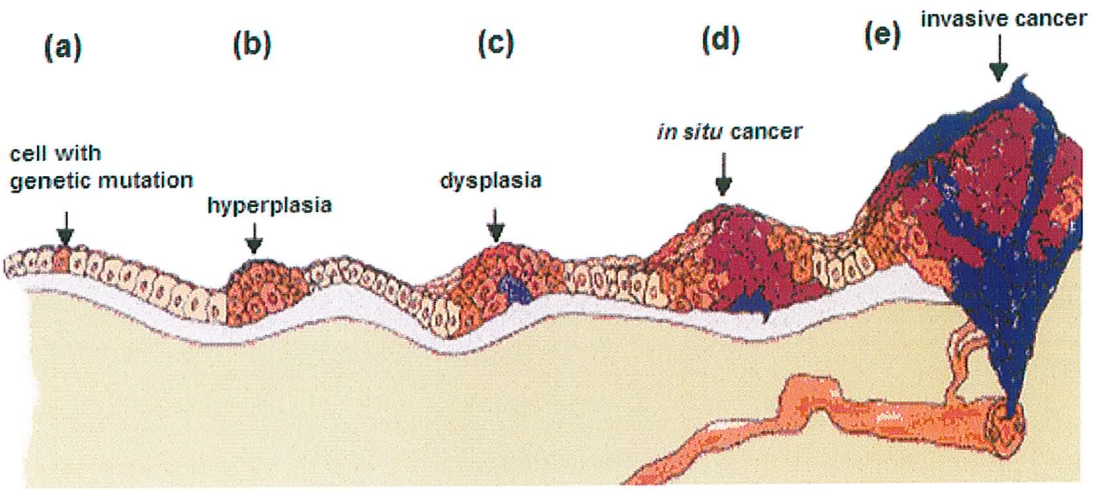


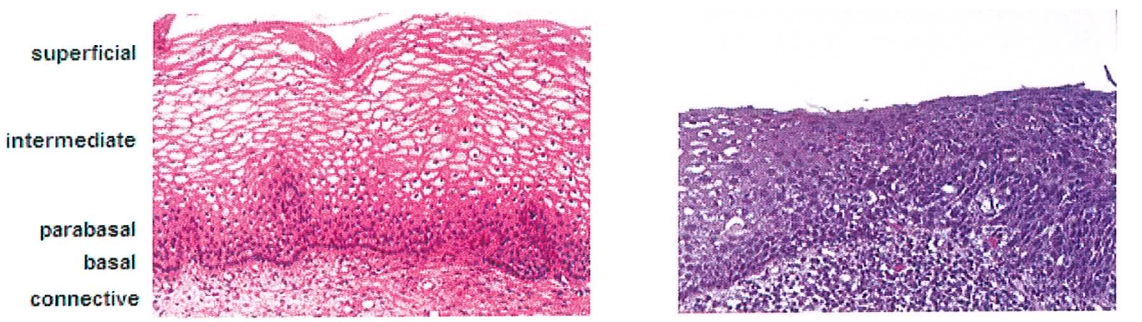
Figure 1.4 The stages of tumour development

(a) genetic mutation, (b) hyperplasia, (c) dysplasia, (d) *in situ* cancer, (e) malignancy

(London Science Museum 2005)

If cells break away from such a tumour, they can travel through the blood stream or the lymph system to other areas of the body and establish new tumours. They continue to grow in new locations. The spread of a tumour to a new site is called *metastasis*. These tumours may become lethal by disrupting the function of a vital organ.

Tumours are of two basic types: benign and malignant. By definition, benign tumours do not invade adjacent tissue orders nor do they metastasise to other sites. The primary descriptor of any tumour is its cell or tissue of origin. Most human malignancies arise from epithelial tissue and are called carcinomas. The "gold standard" in most cancer diagnostics is microscopic evaluation, by a pathologist, of a stained tissue obtained from biopsy of a particular organ (this procedure is called pathologic histology or histopathology). Figure 1.5 shows a comparison between normal cervical structure and dysplastic growth.



(a) **(b)**

Figure 1.5 Photomicrograph of a) normal cervical epithelial cells and b) cervical dysplasia

Figure 1.5 (a) also shows the different regions associated with normal cervical epithelium, namely, basal cells, parabasal cells, intermediate and superficial squamous epithelium, as well as the underlying connective tissue.

The effect of the uncontrolled cell proliferation is evident in Figure 1.5 (b), as the natural structure has been overrun with cancer cells. Although Figure 1.5 shows cervical tissue, malignancies from different tissue types display many of the same characteristics and include the following:

- The morphology and size of cells are different and more variable than those of normal cells.
- The nucleus of a cell that is larger than the nucleus of a normal cell.
- The appearance of large cells with multiple nuclei.
- The invasion of normal tissue by a neoplasm.

The diagnosis of disease and tumour staging is thus in the hands of a pathologist. This diagnosis will lead to a decision by an oncologist regarding the type of cancer treatment to be employed. Although many pathologists are exceptionally good at what they do, this analysis is somewhat subjective. "Misdiagnosis", with a false-negative rate of up to 50% (however more typically 10 – 30%), is especially common in screening of cervical cancer, such as evaluation of Papanicolaou smears (Pap smear) (Wilbur 1997). Additionally, in some cases, a pathologic examination may not produce a firm diagnosis, either because certain tumours are histologically similar or because cells are so poorly differentiated that their tissue of origin cannot be determined.

A diagnostic technique that could replace human interpretation, with a quantitative screening method to detect cancer has the potential to reduce these false-negative rates. This study aims to explore the ability of Raman spectroscopy and Fourier Transform Infrared (FTIR) spectroscopy to fulfill this diagnostic role. Due to the prevalence of cervical cancer among Irish women and the misdiagnosis rate outlined above, this investigation will primarily focus on the potential of vibrational spectroscopy in the diagnosis of cervical cancer.

1.2 Biochemistry – The Building Blocks

Proteins are the most abundant cellular component. They include enzymes, antibodies, hormones and even the cytoskeleton of the cell itself. Structurally and functionally they are the most diverse and dynamic of molecules and play a key role in nearly every biological process. The pattern by which each protein is tailored resides within the genetic information of cells, encoded in a specific sequence of nucleotide bases in DNA. Each such segment of encoded information defines a gene and the expression of the gene leads to the synthesis of the specific protein encoded by it, endowing the cell with the functional attributes unique to that particular protein. This abundance of proteins (constituting more than 50% of the dry weight of cells) and the important role they play makes them a crucial component of cells and worthy of specific spectroscopic investigation.

1.2.1 Amino Acids

Proteins are the indispensable agents of biological function and amino acids are the building block of proteins. Every protein molecule can be viewed as a polymer of amino acids. There are twenty common amino acids. Figure 1.6 shows the generic structure of a single amino acid. At the centre is a tetrahedral carbon atom called the alpha (α) carbon (C_α). It is covalently bonded on one side to an amino group (NH_2) and on the other side to a carboxyl group (COOH). A third bond is always to hydrogen and the fourth bond is to a variable side chain (R).

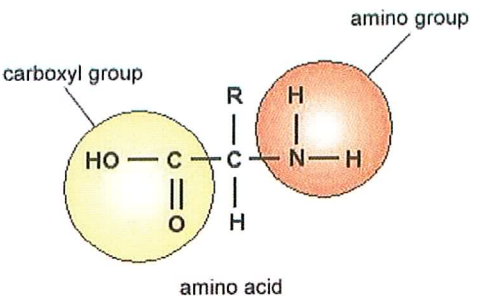


Figure 1.6 A generic amino acid (Kaiser 2005)

Amino acids are grouped according to their side chain properties into 4 classes, namely;

(i) Nonpolar/ hydrophobic amino acids

(ii) Polar/ hydrophilic amino acids

(iii) Acidic amino acids

(iv) Basic amino acids

(i) Hydrophobic (nonpolar) amino acids:

The nonpolar amino acids include all those with alkyl chain R groups (alanine, valine, leucine and isoleucine), as well as proline (with its unusual cyclic structure), methionine (one of the two sulphur containing amino acids) and two aromatic amino acids, phenylalanine and tryptophan (Figure 1.7).

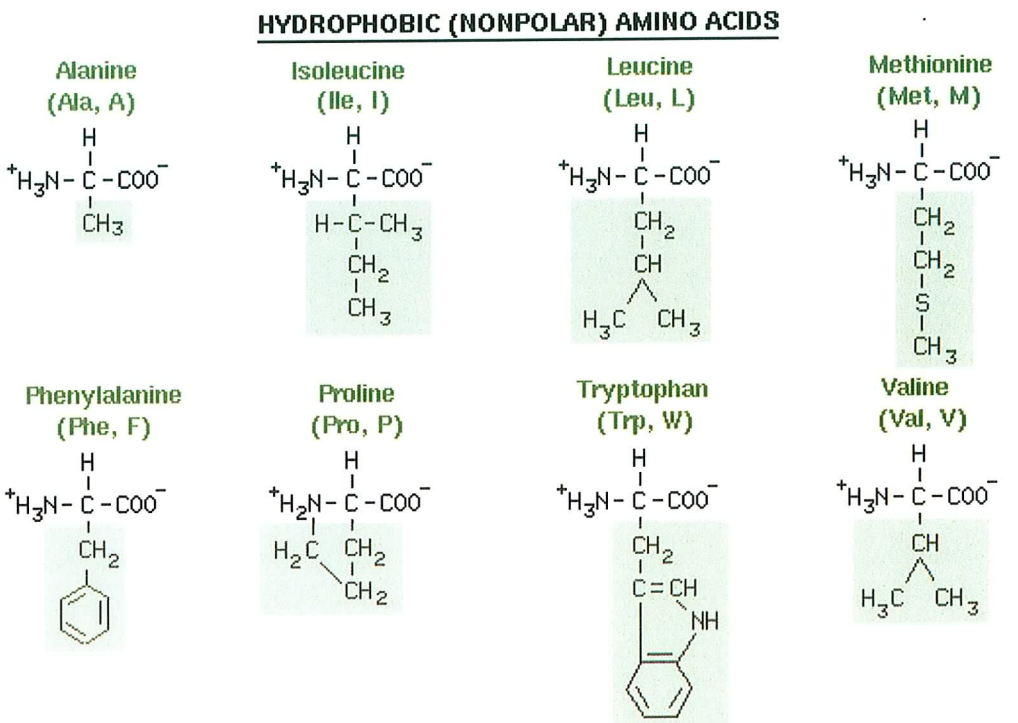


Figure 1.7 Hydrophobic amino acids

(ii) Hydrophilic (polar) amino acids:

The polar, uncharged amino acids, except for glycine, contain R groups that can form hydrogen bonds with water, namely; serine, threonine, cysteine, tyrosine, glycine, asparagine and glutamine. Thus, these amino acids are usually more soluble in water than nonpolar amino acids (Figure 1.8).

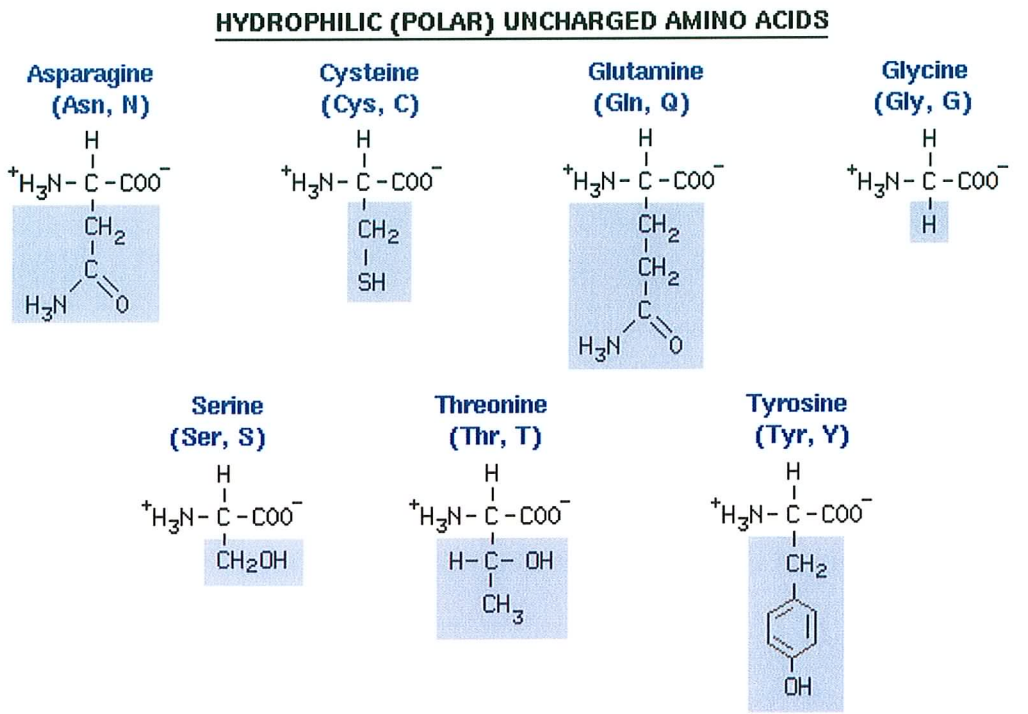


Figure 1.8 Hydrophilic amino acids

(iii) Acidic amino acids:

There are two acidic amino acids - aspartic acid and glutamic acid – whose R groups contain a carboxyl group. These side chain carboxyl groups are weaker acids than the α-COOH group, but are sufficiently acidic to exist as $-COO^-$ at neutral pH (Figure 1.9).

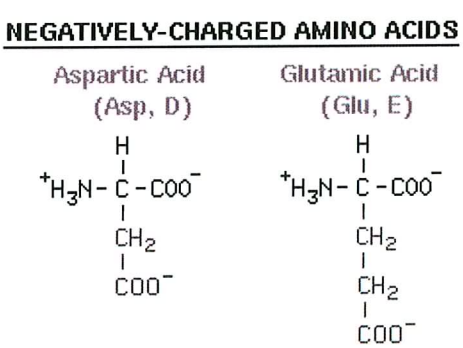


Figure 1.9 Acidic amino acids

(iv) Basic Amino Acids:

Three of the common amino acids have side chains with net positive charges at neutral pH: histidine, arginine and lysine. The ionised group of histidine is an imidazolium, while that of arginine is a guanidinium and lysine contains a protonated alkyl amino group (Figure 1.10).

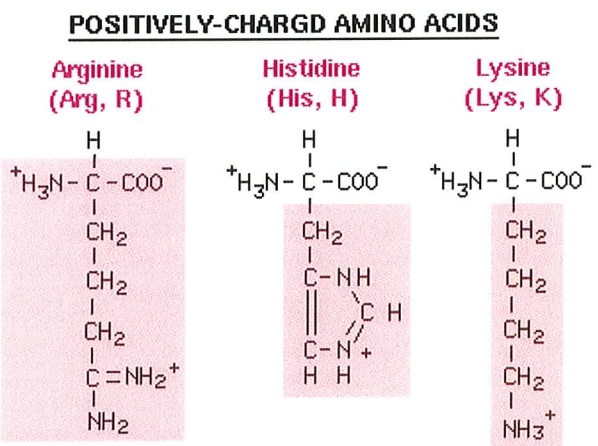


Figure 1.10 Basic amino acids

Amino acids can link together by a covalent peptide bond between the α-carboxyl end of one amino acid and the α-amino end of another. Formally, this bond is formed by the loss of a water molecule as shown in Figure 1.11.

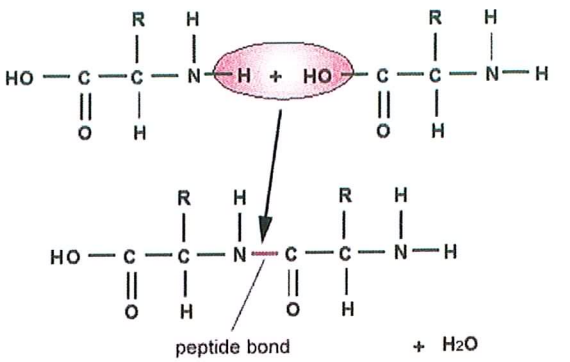


Figure 1.11 Formation of a peptide bond between two amino acids (Kaiser 2005)

The peptide bond has a partial double bond characteristic owing to resonance effects. As a result, the C-N peptide linkage and all of the atoms directly connected to these C and N atoms lie in a planar configuration called the amide plane. This amide plane limits the number of orientations available to the polypeptide chain and plays a major role in determining the three-dimensional structures of proteins.

Any number of amino acids can be joined by successive peptide linkages, forming a polypeptide chain. The polypeptide chain, like the dipeptide, has a directional sense. One end, called the N-terminal or amino-terminal, has a free α-amino group, whereas the other end, the C-terminal or carboxyl-terminal, has a free α-carboxyl group. The sequence of main chain atoms from the N-terminal end to the C-terminal end is C_α–C–N –C_α, etc. and in the opposite direction it is C_α–N–C–C_α, etc. The geometry of the peptide backbone (Fig. 1.11) shows the carbonyl oxygen and the amide hydrogen are *trans* to each other, a conformation favoured energetically due to the reduction in steric hindrance between non-bonded atoms in neighbouring amino acids.

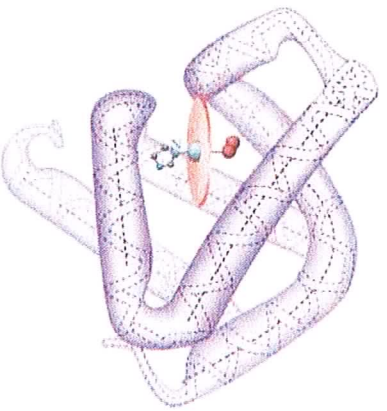
Short polypeptide chains, up to a length of about 20 amino acids are called peptides and fragments of whole polypeptide chains are called oligopeptides. A small protein molecule may contain a polypeptide chain of only 50 amino acids. A large protein may contain chains of 3,000 or more amino acids.

1.2.2 Proteins

The conformation of a native or highly organised protein reflects a delicate balance among a variety of interaction forces, both within the protein's interior and with the surrounding solvent. If solvent surrounding the proteins is perturbed, the protein's native conformation can be disrupted, with a resulting loss of function and the production of a partially unfolded, or denatured, protein.

As a first approximation, proteins can be assigned to one of two classes on the basis of shape: fibrous or globular (Figure 1.12).

(A) Collagen, a fibrous protein



(B) Myoglobin, a globular protein

Figure 1.12 Two classes of protein shape, (A) fibrous and (B) globular (Mallery 2005)

Fibrous proteins tend to have relatively simple, regular line structures. These proteins often serve structural roles in cells. Typically, they are insoluble in water or in dilute salt solutions. In contrast, globular proteins, are roughly spherical in shape. The polypeptide chain is compactly folded so that its hydrophobic amino acid side chains are in the interior of the molecule and the hydrophilic side chains are on the outside exposed to the solvent, water. Consequently, globular proteins are usually very soluble in aqueous solutions. Most soluble proteins of the cell, such as the cytosolic enzymes, are globular in shape.

The architecture of protein molecules is quite complex. Nevertheless, this complexity can be resolved by defining various levels of structural organisation.

Primary Structure

The amino acid sequence is the primary (1°) structure of a protein, as shown in Figure 1.13, for example.

Figure 1.13 Example of the primary structure of a protein (Kaiser 2005)

Secondary Structure

Through hydrogen bonding interactions between adjacent amino acid residues, the polypeptide chain can arrange itself into characteristic patterns, such as pleated sheet or helical segments. These segments constitute structural conformities, so-called regular structures that extend along one dimension, like the coils of a spring. Such architectural features of a protein are designated as secondary (2°) structures (Fig. 1.14).

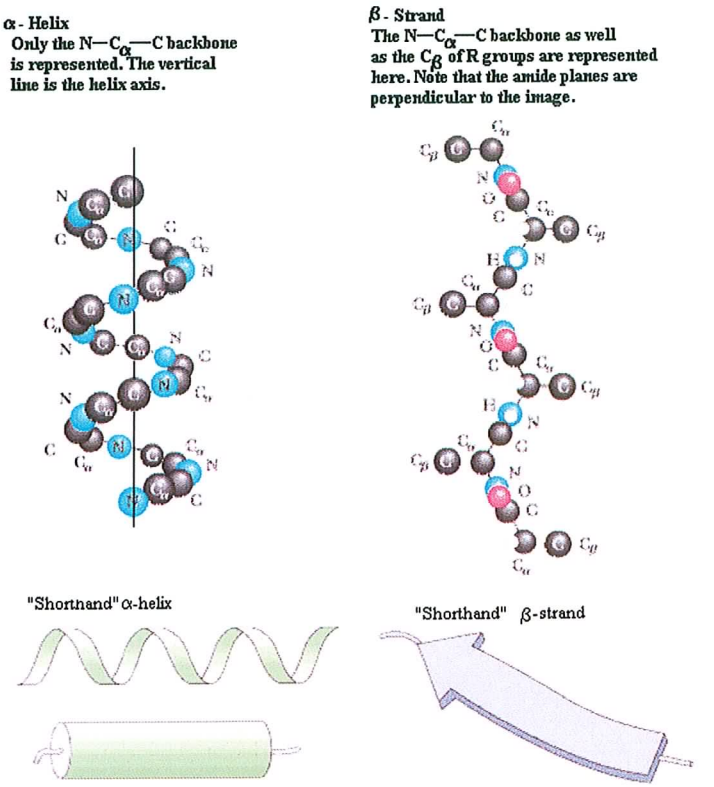


Figure 1.14 Two structural motifs that arrange the primary structure of proteins into a higher level of organization (Harcourt 1999)

Tertiary Structure

When the polypeptide chains of protein molecules bend and fold in order to assume a more compact three-dimensional shape, a tertiary (3°) level of structure is generated (Figure 1.15). This is a result of interactions between amino acid side groups and includes;

- van der Waals bonds (if the side groups are nonpolar)
- Hydrogen bonds (if the side groups contain hydroxyl or amino groups)

- Ionic bonds (if the side groups are acids and bases that can transfer protons from one to another making a carboxylate ion, which is negative and essentially an ammonium or quaternary ammonium ion which is positive)
- Covalent/ disulfide bonds (if the side groups are cysteine residues in which the sulphur atoms are bonded together by the removal of two hydrogen atoms)
- Hydrophobic interactions

It is by virtue of their tertiary structure that proteins adopt a globular shape, especially those proteins commonly existing in solution in the aqueous compartments of cells. A globular conformation gives the lowest surface-to-volume ratio, shielding much of the protein from interaction with the solvent.

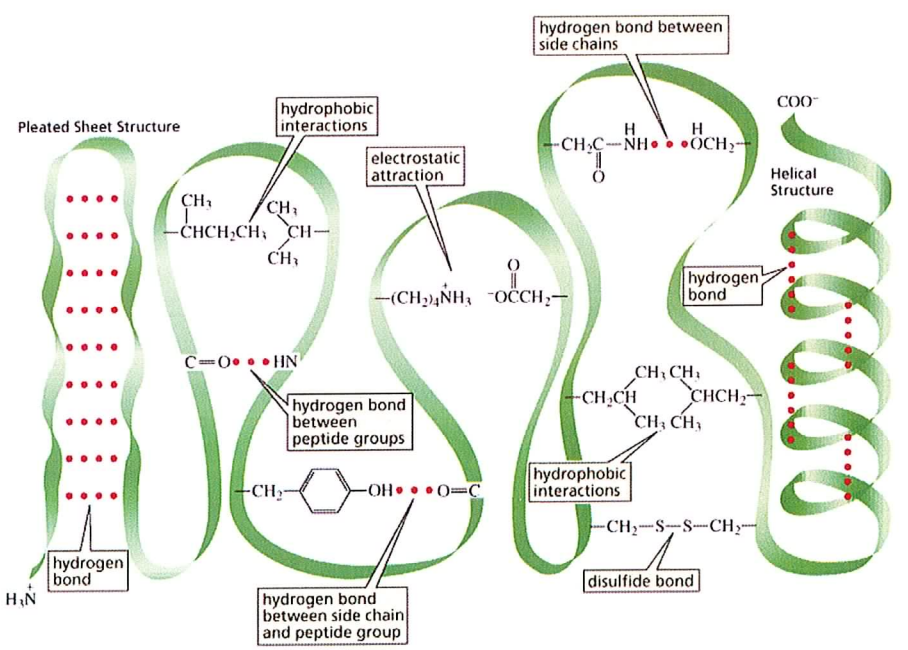


Figure 1.15 Interactions resulting in tertiary structure of proteins (Harcourt 1999)

Quaternary Structure

Many proteins consist of two or more interacting polypeptide chains of characteristic tertiary structure, each of which is commonly referred to as a subunit of the protein. Subunit organisation in proteins constitutes another level in the hierarchy of protein structure, defined as the protein's quaternary (4°) structure (Fig 1.16). Figure 1.17 summarises all four levels of protein structure.

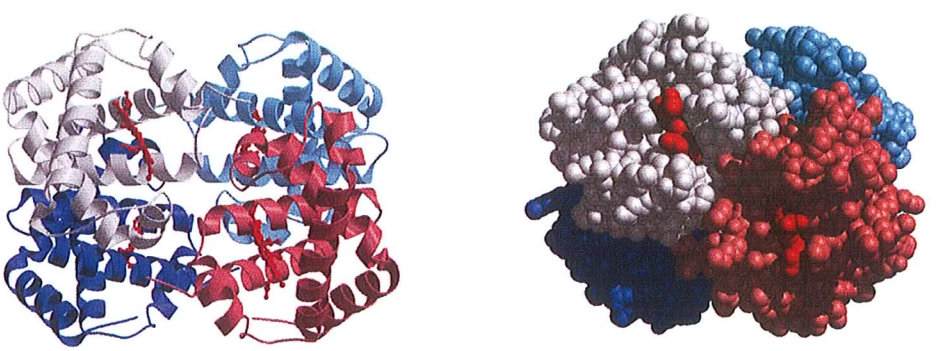


Figure 1.16 Quaternary structure of a protein with 4 subunits (Harcourt 1999)

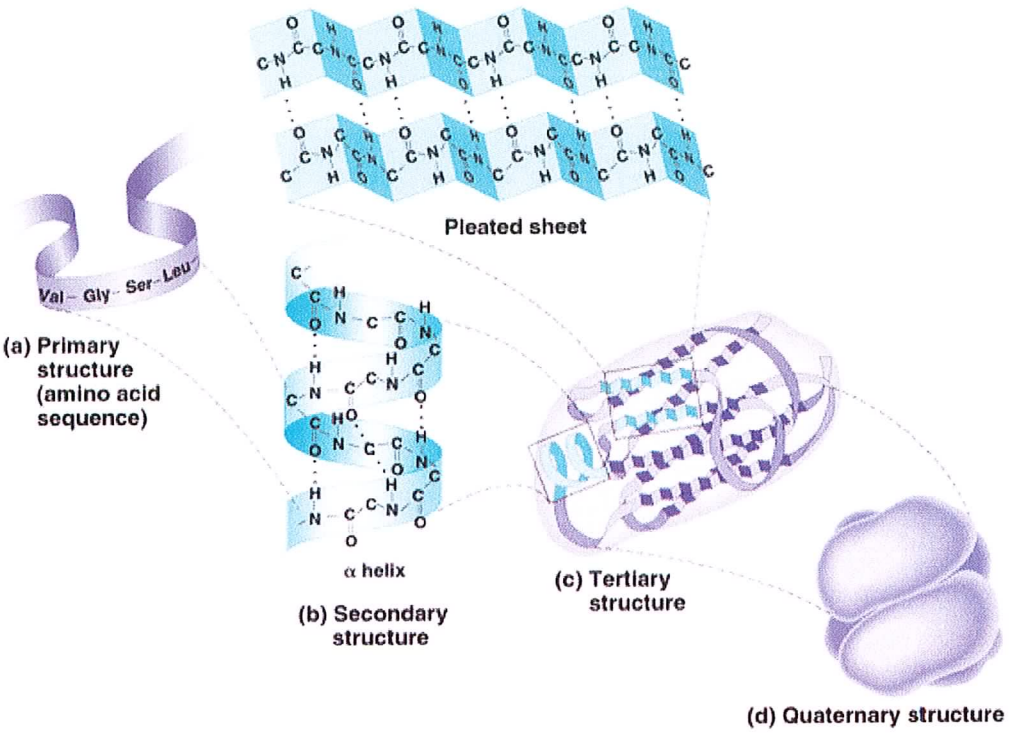


Figure 1.17 A summary of all the four levels of structural organisation in proteins (Harcourt 1999)

1.2.3 Nucleic Acids

Nucleotides and nucleic acids are biological molecules that possess heterocyclic nitrogenous bases as principal components of their structure. The biochemical roles of nucleotides are numerous. They participate as essential intermediates in virtually all aspects of cellular metabolism. Serving an even more central biological purpose are the nucleic acids, the elements of heredity and the agents of genetic information transfer.

Just as proteins are linear polymers of amino acids, nucleic acids are linear polymers of nucleotides. The two basic kinds of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the repository of genetic information in cells, while RNA serves in the transcription and translation of this information.

The bases of nucleotides and nucleic acids are derivatives of either pyrimidine or purine. Pyrimidines are six-membered heterocyclic aromatic rings containing two nitrogen atoms (Fig. 1.18). The purine ring structure consists of a combination of a pyrimidine ring with a five-membered imidazole ring to yield a fused ring system (Fig. 1.19)

The common naturally occurring pyrimidines are cytosine, thymine and uracil. Cytosine and thymine are the pyrimidines found in DNA, whereas cytosine and uracil are common in RNA. Adenine and guanine, the two common purines, are found in both DNA and RNA.

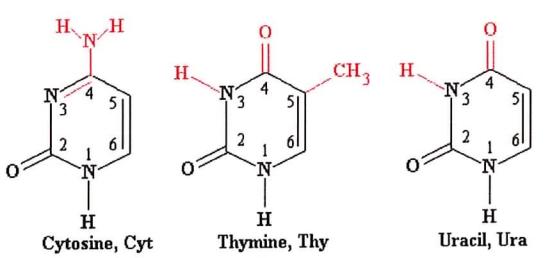


Figure 1.18 Common pyrimidines

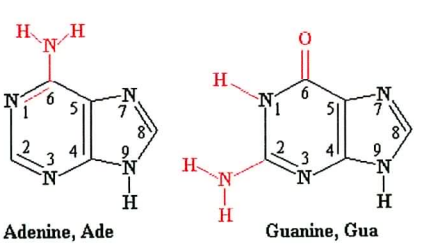


Figure 1.19 Common purines

1.2.4 Lipids

The term lipid refers to a diverse range of molecules and to some extent it is a catch-all term for relatively water-insoluble or nonpolar compounds of biological origin, including waxes, fatty acids, fatty acid-derived phospholipids, sphingolipids, glycolipids and terpenoids such as retinoids and steroids. Some lipids are linear aliphatic molecules, while others have ring structures.

Most lipids have some polar character in addition to being largely nonpolar. Generally, the bulk of their structure is nonpolar or hydrophobic, meaning that it does not interact well with polar solvents like water. Another part of their structure is polar or hydrophilic and will tend to associate with polar solvents like water. This makes them amphiphilic molecules (having both hydrophobic and hydrophilic portions). In the case of cholesterol, the polar group is a mere -OH (hydroxyl or alcohol). In the case of phospholipids, the polar groups are considerably larger and more polar.

Phospholipids or, more precisely glycerophospholipids, are built on a glycerol core to which are linked two fatty acid-derived "tails" by ester linkages and one "head" group by a phosphate ester linkage (Fig. 1.20). Fatty acids are unbranched hydrocarbon chains, connected by single bonds alone (saturated fatty acids) or by both single and double bonds (unsaturated fatty acids). The chains are usually 10-24 carbon groups long. The head groups of the phospholipids found in biological membranes are phosphatidylcholine (lecithin), phosphatidylethanolamine, phosphatidylserine and phosphatidylinositol, whose head group can be modified by the addition of one to three more phosphate groups. While

phospholipids are the major component of biological membranes, other lipid components like sphingolipids and sterols (such as cholesterol in animal cell membranes) are also found in biological membranes.

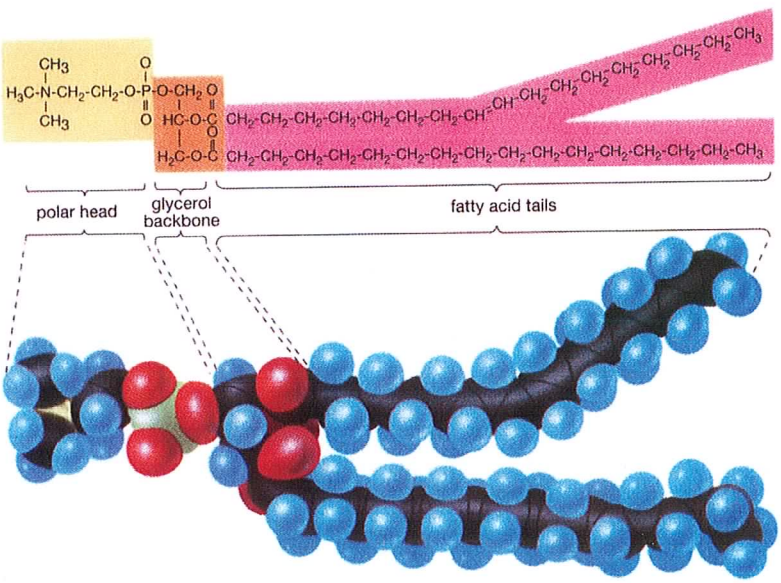


Figure 1.20 Molecular structure of the phospholipid lecithin (Sinauer 1998)

1.2.5 Carbohydrates

All carbohydrates are made up of units of sugar (called saccharide units). Carbohydrates that contain only one sugar unit (monosaccharides) or two sugar units (disaccharides) are referred to as simple sugars. Simple sugars are broken down quickly in the body to release energy. Two of the most common monosaccharides are glucose and fructose. Glucose ($C_6H_{12}O_6$) is the primary form of sugar stored in the human body for energy (Fig. 1.21).

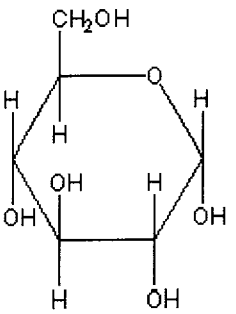


Figure 1.21 Molecular structure of glucose (Stryer 2005)

Complex carbohydrates (polysaccharides) are polymers of the simple sugars. Starch is the principal polysaccharide used by plants to store glucose for later use as energy. When starch is eaten, amylase (an enzyme that occurs in saliva and in the intestines) breaks the bonds between the repeating glucose units thus allowing the sugar to be absorbed into the bloodstream. Once absorbed into the bloodstream, glucose is distributed to the areas where it is needed for energy or stored as the polymer / polysaccharide glycogen (Fig. 1.22). Excess glucose is bonded together to form glycogen molecules.

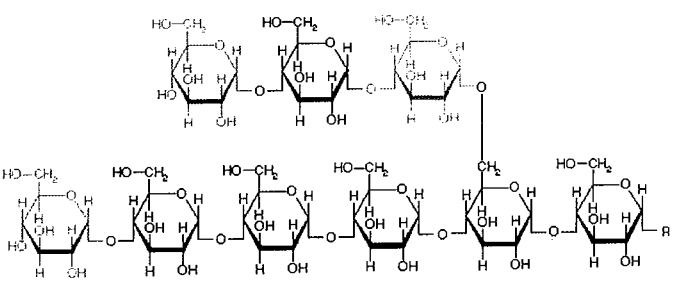


Figure 1.22 Molecular structure of glycogen (Stryer 2005)

1.3 Vibrational Spectroscopy

Vibrational spectroscopy provides useful information on the structure of molecules. Molecules absorb radiation if the frequency of the radiation absorbed is exactly the same as the vibrational frequency of the molecule. These frequencies are dependent on:

1. The strength of the bond between atoms
2. The atomic weight of the atoms
3. The molecule's geometry; its bond angles and bond lengths

The two main techniques involved in the analysis of the vibrational states of molecules are infrared (IR) and Raman spectroscopy.

1.3.1 Infrared Spectroscopy

The term infrared covers the range of the electromagnetic spectrum between 0.78 and 1000 μm. In the context of infrared spectroscopy, the range is usually measured in wavenumbers. The infrared region of the spectrum is divided into three sections according to wavelength (or wavenumber); near, mid and far infrared (Table 1.1)

Region	Wavelength range (μm)	Wavenumber range (cm^{-1})
Near	0.78 - 2.5	12800 - 4000
Middle	2.5 – 50	4000 - 200
Far	50 –1000	200 – 10

Table 1.1 Infrared spectral regions

IR radiation does not have enough energy to induce electronic transitions. Absorption of IR radiation is restricted to compounds with small energy differences in the possible vibrational and rotational states. For a molecule to absorb IR radiation, the vibrations or rotations within a molecule must cause a net change in the dipole moment of the molecule. The alternating electrical field of the radiation interacts with fluctuations in the dipole moment of the molecule. If the frequency of the radiation matches the vibrational frequency of the molecule then radiation will be absorbed, causing a change in the amplitude of molecular vibrations. The positions of atoms in molecules are not fixed and are subject to a number of different vibrations. Vibrations fall into the two main categories of stretching (change in inter-atomic distance along the bond axis) and bending (change in angle between two bonds). The motions involved in CO stretching and CCH bending, for example, are illustrated in Figure 1.23.

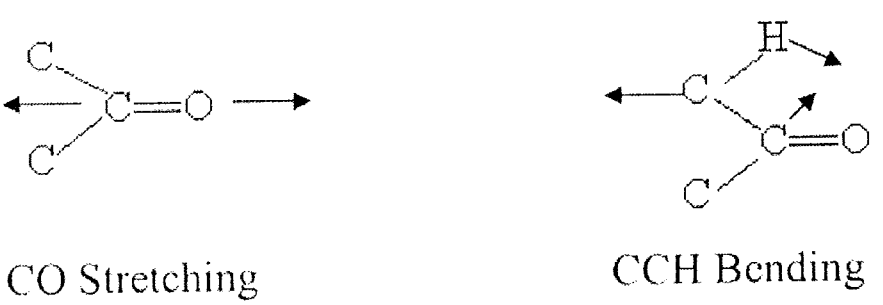


Figure 1.23 CO stretching and CCH bending modes

In order to understand the IR spectrum of a molecule the fundamental frequencies and symmetry must be considered. A molecule consisting of N atoms has a total of 3N degrees of freedom, corresponding to the Cartesian coordinates of x, y and z of each atom of the molecule. In a non-linear molecule 3 of these degrees are rotational and 3 are translational,

giving 3N-6 normal modes of vibration. In a linear molecule 2 degrees are rotational and 3 are translational giving 3N-5 normal modes of vibration.

The vibrational motion of a diatomic molecule can be approximated by classical mechanics. In this approximation two atoms and a connecting bond are treated as a simple harmonic oscillator composed of two masses (m_1 and m_2) joined by a spring. If a displacement of the two atoms (Δr) is carried out there will be a force acting to return the spring to the equilibrium position. According to Hooke's law the restoring force exerted by the spring (f) is proportional to the displacement Δr according to equation 1.1.

$$F = -k\Delta r \qquad \textit{Equation 1.1}$$

where k (Hooke's law constant for the spring) is called the force constant for a molecular system.

For harmonic oscillation of a diatomic molecule the potential energy, E, is given by:

$$E = \frac{1}{2}kx^2 \qquad \textit{Equation 1.2}$$

where x is the displacement of the two masses from the equilibrium position. A plot of potential energy as a function of x is thus a parabola that is symmetrical about the equilibrium internuclear distance as a minimum (Figure 1.24).

However this classical spring model does not hold for molecules because their vibrational motion is quantized and a quantum mechanical treatment of molecular systems is given in equation 1.3

$$E = \left(n + \frac{1}{2}\right)h\nu \qquad \textit{Equation 1.3}$$

where υ is the frequency of the vibration, h is Planck's constant and n is the vibrational quantum number (0, 1, 2, 3….). The addition of the ½ hν accounts for the zero point energy of the system. These states are indicated for a harmonic oscillator in Figure 1.24.

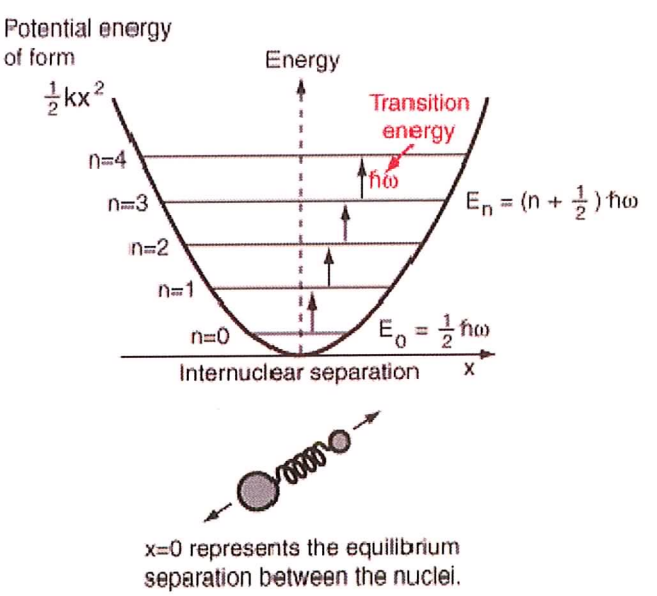


Figure 1.24 Potential Energy versus distance x for a harmonic oscillator showing the vibrational states

The potential energy curve of a real molecule is not a perfect parabola since the vibrational energy levels are not equally spaced but converge because the molecule undergoes anharmonic oscillation. However the assumption of harmonic oscillation can be sufficient for certain purposes such as describing fundamental vibrations.

The difference in energy in wavenumbers, between two adjacent levels is given by equation 1.4

$$\bar{v} = \frac{1}{2\pi c}\sqrt{\frac{k(m_1 + m_2)}{m_1 m_2}} \qquad \textit{Equation 1.4}$$

where $\bar{v}$ is the vibrational frequency (cm^{-1}), m_1 and m_2 are the masses of atoms 1 and 2 in grams, c is the velocity of light in (cm/s) and k is the force constant of the bond (dyne/cm). This equation shows the relationship of bond strength and atomic mass to the wavenumber at which the molecule will absorb. Therefore as the mass of an atom decreases the vibration frequency increases. Similarly a stronger double bond results in a higher vibrational frequency.

The symmetry of the molecule plays an important role in determining whether a vibration is strongly observable through infrared absorption spectroscopy. As with other electrical dipole transitions, a change in the equilibrium dipole moment must be affected for the transition to be allowed. The symmetric stretch in linear molecules such as CO_2 is inactive in the IR because this vibration produces no change in the dipole moment of the molecule.

The IR active vibrations of moieties within a molecule are represented in the spectrum by identifiable bands. Thus the spectrum of a molecule can be used as a fingerprint by which the molecule can be identified as each vibrational mode has a characteristic frequency.

1.3.2 FTIR Instrumentation

IR absorption measurements in this work were carried out using a Perkin Elmer Spectrum GX which is a single-beam, Michelson interferometer based, Fourier Transform infrared spectrometer (Fig. 1.25). It has a dual level optical module that is sealed and desiccated. The system is configured with a mid-infrared single source. MIR and FIR beam splitters and DTGS detector kits allow the range 7000 to 50cm^{-1} to be covered with a maximum resolution of 0.3cm^{-1}. The Spectrum GX is a modular system and can accommodate up to four equivalent output beams. The spectrometer is configured with the Auto IMAGE microscope system, which can operate in transmission or reflectance modes. All microscope operations including adjustments to aperture, focus and illumination are fully automated and controlled from the PC. It includes a built-in 35W tungsten halogen illuminator, a motorised stage and a CCD video camera. The medium beam MCT detector covers the range from 5500 to 550 cm^{-1}. An ATR attachment with a micro germanium crystal with a range from 5500 to 600 cm^{-1} can be used for micro samples and ATR mapping for surface studies.

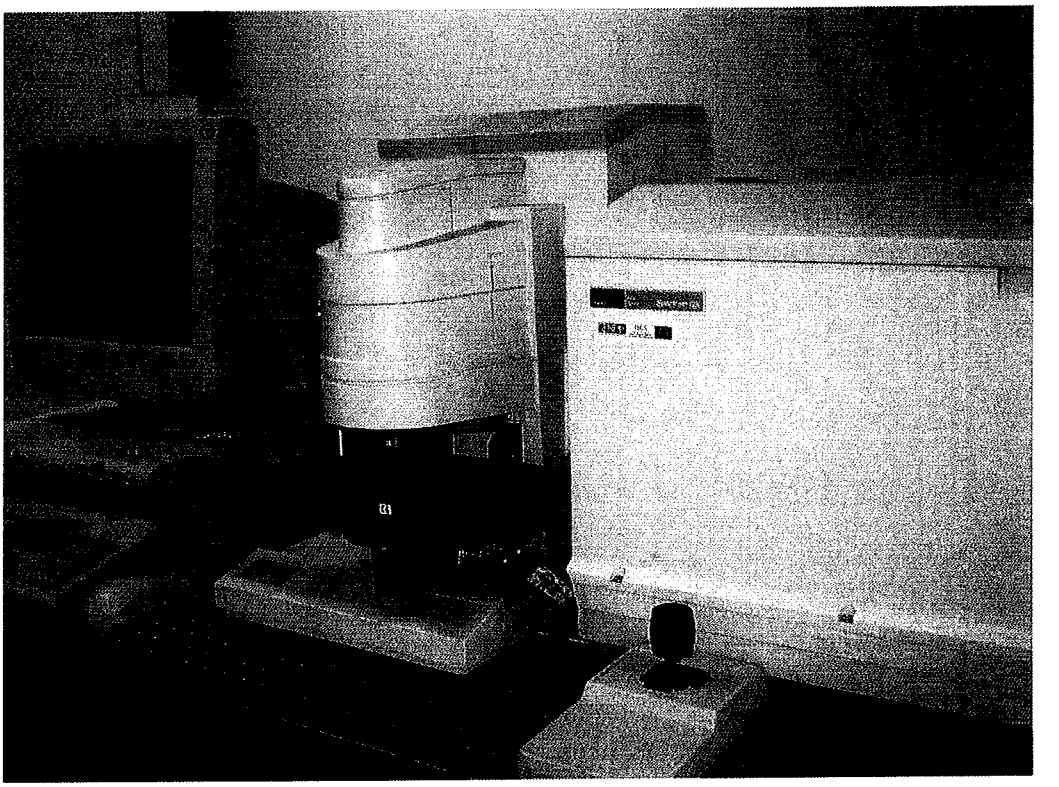


Figure 1.25 Perkin Elmer Spectrum GX spectrometer

Synchrotron FTIR measurements were taken at the 13.3 beamline at the synchrotron radiation source (SRS), Daresbury Laboratory, Cheshire, England. The beam is extracted from a storage ring of current ≈250mA at ≈2GeV. The station utilises an FTIR microscope (Nicolet Nexus FTIR spectrometer and a Nicolet Continuμm IR microscope, ThermoNicolet Inc., Madison, USA) interfaced to the beamline exit port. The beamline is deflected using a series of mirrors, including a cooled mirror in the plane of the ring, so as to enter the spectrometer (Bozec *et al.* 2002). A Spectra-Tech 32X Reflachromat objective (0.65 N.A. and 8 mm working distance) was used (Fig. 1.26).

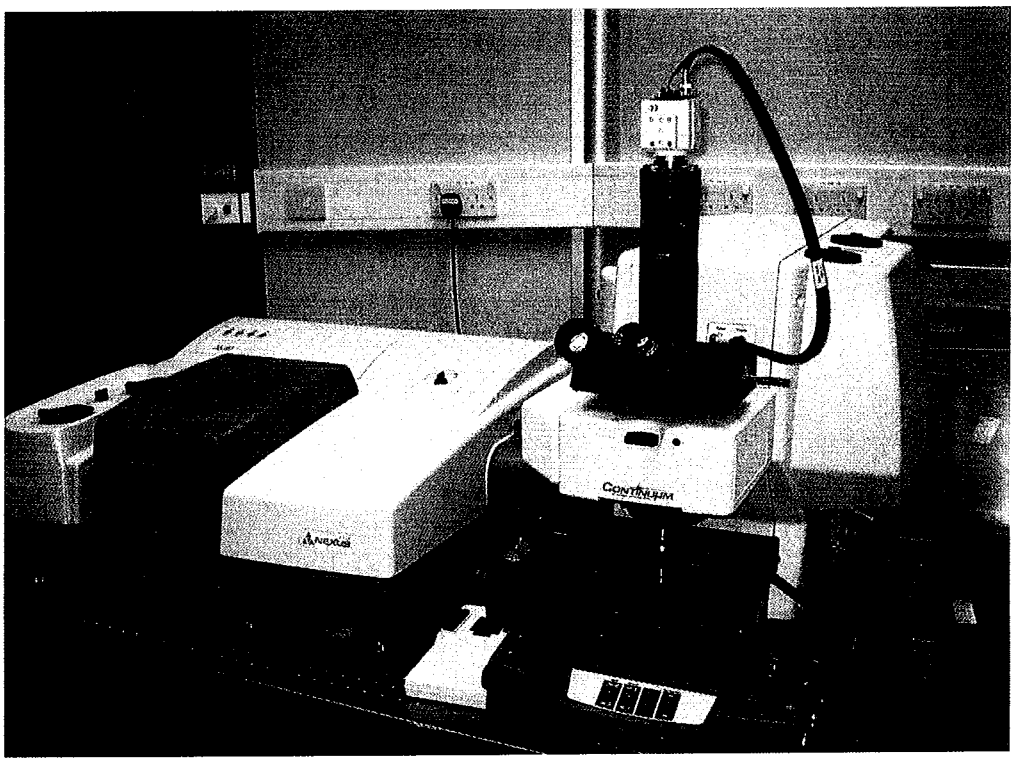


Figure 1.26 Nicolet Nexus FTIR spectrometer

1.3.3 Raman Spectroscopy

Rayleigh scattering (elastic scattering) occurs when the scattered light is of the same frequency as the incident light. Raman scattering (inelastic scattering) is a result of light that is scattered off a molecule where its frequency differs from that of the incident light as a result of the interaction of the photon with the material vibrations. In Raman scattering the energy increase or decrease from the excitation is related to the vibrational energy spacing in the ground electronic state of the molecule and therefore the wavenumber shift of the Stokes and anti-Stokes lines are a direct measure of the vibrational energies of the molecule. In Stokes Raman scattering, the molecule starts out in a lower vibrational energy

state and after the scattering process ends up in a higher vibrational energy state. Thus the interaction of the incident light with the molecule creates a vibration in the material. In anti-Stokes scattering, the molecule begins in a higher vibrational energy state and after the scattering process ends up in a lower vibrational energy state. Thus a vibration in the material is annihilated as a result of the interaction. The frequency differences between the Raman lines and the exciting line are characteristic of the scattering substance and are independent of the frequency of excitation. The Raman effect arises from the coupling of the induced polarisation of scattering molecules that is caused by the electric vector of the electromagnetic radiation with the molecular vibrational modes. Light of frequency, ω_L produces a polarisation in a material given by equation 1.5.

$$P(\omega_L) = \chi(\omega_L)\mathrm{E}_0 Cos\omega_L t \qquad \textit{Equation 1.5}$$

where P is the polarisation, ω_L is the frequency of incident light, E is the electric field and $\chi(\omega_L)$ is the polarisability or susceptibility, considered to be a constant of the material associated with its electronic properties. However at a finite temperature a material is not at equilibrium and atoms will vibrate about an equilibrium position, R, along the normal co-ordinates, with frequency ω_K in accordance with a simple harmonic oscillator approximation. The displacement from equilibrium can be represented by equation 1.6;

$$\Delta R(t) = \Delta R Cos(\omega_K t) \qquad \textit{Equation 1.6}$$

The susceptibility to polarisation thus varies through a cycle and can be represented by equation 1.7;

$$\chi_k(t) = \chi_0 + \Delta\chi_k Cos(\omega_k t) \qquad \textit{Equation 1.7}$$

The polarisation now has the form as illustrated in equation 1.8;

$$P(\omega_L, \omega_K) = \chi_0(\omega_L)E_0 Cos\omega_L t + \Delta\chi_K E_0 Cos(\omega_L)tCos(\omega_K - \delta_K) \qquad \textit{Equation 1.8}$$

where δ_K takes into account any phase difference between the molecular vibration and the electric field oscillation. This may be written as Equation 1.9:

$$P(\omega_L, \omega_K) = \chi_0(\omega_L)E_0 Cos\omega_L t + 1/2\Delta\chi_K E_0 (Cos((\omega_L - \omega_K)t - \delta_K) + Cos((\omega_L + \omega_K)t + \delta_K))$$

Equation 1.9

Thus the polarisation has the form:

$$P = P(\omega_0) + P(\omega_0 - \omega_K) + P(\omega_0 + \omega_K) \qquad \textit{Equation 1.10}$$

An oscillating polarisation will reradiate at the oscillation frequency and thus the scattered light has three components. $P(\omega_0)$ corresponds to Raleigh scattering. $P(\omega_0 - \omega_K)$ corresponds to the subtraction of a vibrational quantum from the photon energy resulting in the creation of a vibration – termed the Stokes lines of the Raman spectrum. $P(\omega_0 + \omega_K)$ corresponds to the addition of a vibrational quantum to the photon by the annihilation of a

vibration - termed the anti-Stokes lines of the Raman spectrum. The ratio of the intensity of Raman Stokes and anti-Stokes lines is given in equation 1.11;

$$\frac{I_S}{I_{AS}} = \left(\frac{\omega_0 + \omega_K}{\omega_0 - \omega_K}\right)^4 e^{\left(\frac{h\nu}{kT}\right)} \qquad \textit{Equation 1.11}$$

In Raman spectroscopy, each vibration can couple to the laser generating a vibrational spectrum on both the Stokes and anti-Stokes sides. The Stokes shift is normally measured, as at room temperatures it is easier to create a phonon than to annihilate one, in accordance with the Boltzmann distribution. The ratio between the Stokes and anti Stokes scattered intensities, as represented in equation 1.11, can be used as a measure of temperature.

In Raman spectroscopy, the parameter of interest is the frequency shift as a result of the coupling to the vibrations and thus the incident frequency is set to zero, the Stokes line being represented as a positive shift. For ease of comparison to infrared spectroscopy, frequency shifts are expressed in wavenumbers.

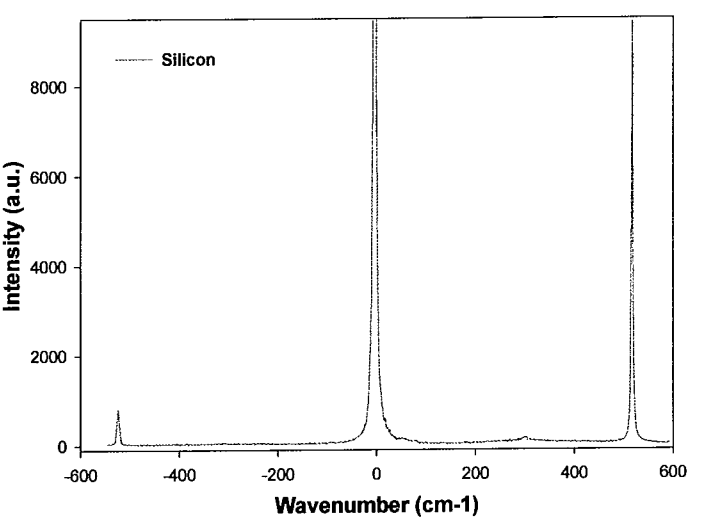


Figure 1.27 Raman spectrum of a silicon crystal

Figure 1.27 shows the Raman spectrum of a Silicon crystal. The Stokes (positive) and anti-Stokes (negative) Raman lines can be seen symmetrically shifted from the incident laser line.

The selection rule for Raman activity states that there must be a change in the polarisability of a molecule during the vibration. For example, the symmetric stretch of CO_2, although IR inactive is Raman active and the band occurs at $1351cm^{-1}$ in its Raman spectrum. Therefore, although Raman spectroscopy is very similar to the more frequently used IR spectroscopic technique, the two vibrational spectroscopic techniques are, in fact, very complementary. For polar molecules, some vibrations give rise to strong bands in the IR spectrum since they cause large changes in the dipole moment, but they give weak Raman signals since it is difficult to change the polarisability of such molecules. Like-wise, for non-polar molecules, vibrations that give very strong Raman bands usually result in weak infrared signals. For example, hydroxyl or amine stretching vibrations and the vibrations of carbonyl groups, are usually very strong in an IR spectrum and usually weak in a Raman spectrum. However, the stretching vibrations of carbon double or triple bonds and symmetric vibrations of aromatic groups are very strong in the Raman spectrum. Raman has the advantage of minimal interference from water so it is a good choice for biological samples with a view to in-vivo measurements. Also, as Raman is usually carried out using visible wavelengths, higher spatial resolution can be achieved.

1.3.4 Raman Instrumentation

The Instruments SA Labram system is a confocal Raman imaging microscope system with a motorised XY sample stage for automated Raman imaging (Fig. 1.28). Both Helium-

Neon (632.8nm/11mW) and Argon ion (514.5nm/130mW, 488nm/130mW, 457nm/20mW) lasers are available as sources. Both are polarised, enabling measurement of depolarisation ratios and studies of orientation in materials. The light is imaged to a diffraction limited spot (typically 2μm) via the X50 objective lens of an Olympus microscope (although other lenses are available – X10, X20 and X100). The scattered light is collected by the objective lens in a confocal geometry and is dispersed onto an air cooled CCD detector (1024x256 pixels) by one of two interchangeable gratings, 1800 lines/mm or 600 lines/mm, allowing the range from 150cm^{-1} to 4000cm^{-1} to be covered in a combination of images. The resolution of the system operating with the 1800 lines/mm is 1.65cm^{-1}/pixel. The confocal microscopic system allows measurement of powdered samples with no further sample preparation, direct measurement of solids, liquids and solutions. Spectral X-Y mapping may also be performed. Furthermore the system can be equipped with a remote head (with the use of a fiber optic extension) coupled to the spectrometer.

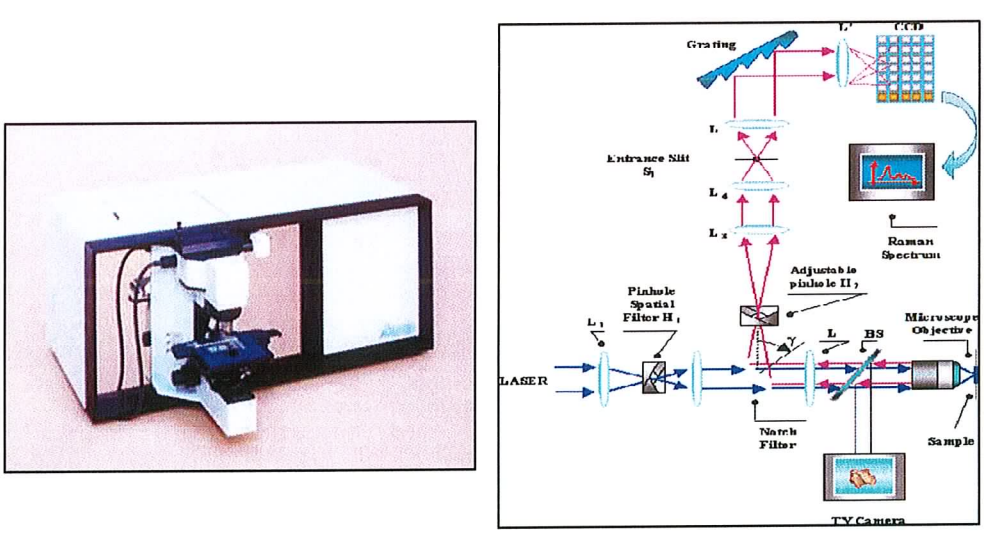


Figure 1.28 Labram Raman spectrometer and schematic layout

1.4 Vibrational spectroscopy for cancer detection

There appears to have been much more activity in the area of FTIR spectroscopy (as opposed to Raman) and the study of cancer. Studies have been carried out on a variety of organs and a selection includes; **cervical** (Wong *et. al* 1991, Wong *et. al* 1993, Wong *et al.* 1995, Wong *et al.* 2002, Yazdi *et al.* 1996, Wood *et al.* 1996, Fung *et al.* 1997, Shaw et al. 1999, Wood *et al.* 1998, Cohenford *et. al* 1997, Cohenford *et al.* 1998, Chiriboga *et al.* 1998, Neviliappan *et al.* 2002, Romeo *et al.* 2002 a,b, Diem *et al.* 2002, Mantsch *et al.* 2002, Chang *et al.* 2003, Wood *et al.* 2003, Mordechai *et al.* 2004, Romeo *et al.* 2004, Diem *et al.* 2004), **brain** (Levin *et al.* 2005, Kneipp *et al.* 2003b, Dovbeshko *et al.* 2000, Keller *et al.* 1994), **breast** (Ling *et al.* 2005, Li *et al.* 2004, Bindig *et al.* 2003, Shah *et al.* 2001), **endometrium** (Sindhuphak *et al.* 2003, Fung *et al.* 1996), **heart** (Wang *et al.* 2005, Severcan *et al.* 2003, Liu *et al.* 1996), **liver** (Chiriboga *et al.* 2000), **lung** (Ren *et al.* 2003), **prostate** (Paluszkiewicz *et al.* 2001), **skin** (McIntosh *et al.* 2002) and **thyroid** (Liu *et al.* 2003).

Although not as frequently used as FTIR spectroscopy, Raman spectroscopy has been employed in many cancer studies, of many different cells/tissues. There are a number of reasons why the use of Raman spectroscopy has not been as widespread as FTIR spectroscopy. Raman spectroscopy traditionally has been used to study the structure and function of molecules by analytical chemists and biochemists. Raman spectroscopy of biological samples is different from that of other types of samples for a number reasons. Biological molecules tend to be large macromolecules and samples tend to be

inhomogeneous mixtures in terms of chemical structure and morphology. Raman scattering is inherently weak. For this reason it is difficult to observe without intense monochromatic excitation and sensitive detectors. It was not until the 1960's with the development of commercial lasers and detectors that the modern era of Raman spectroscopy began. Also, the complex nature of tissue composition results in absorption of light throughout the entire UV-visible region and subsequent fluorescence emission can strongly interfere with weak Raman signals. These problems were addressed in the 1990's with the advent of sensitive detectors, spectrographs and tunable lasers (Manoharan *et al.* 1996). Many groups now use UV or NIR FT Raman spectroscopy for recording fluorescence-free spectra from human tissues.

Raman spectroscopy has been used in the detection of a variety of cancers such as **cervical** (Utzinger *et al.* 2001, Mahadevan-Jansen *et al.* 1998, Yazdi *et al.* 1998, Liu *et al.*1992), **breast** (Kneipp *et al.* 2003a, Smith *et al.* 2003), **skin** (Nunes *et al.* 2003, Nijssen *et al.* 2002, Caspers *et al.* 2001, Gniadecka *et al.* 1997), **lung** (Huang *et al.* 2003, Kaminaka *et al.* 2001), **brain** (Mizuno *et al.* 1994), **bladder** (Crow *et al.* 2003), **oesophageal carcinoma** (Boere *et al.* 2003, Kendall *et al.* 2003, Song *et al.* 2003), **colon** (Molckovsky *et al.* 2003), liver (Hawi *et al.* 1996) and **prostate cancer** (Crow *et al.* 2003).

1.5 Conclusion

With considerably more research having been carried out using IR spectroscopy (as opposed to Raman) of cells and tissues, there is much to be learnt from using the two techniques side by side on the same samples. The results of Raman research in the area of cervical cancer to date have had many contradictory findings, which require further investigation and will be addressed in Chapter 6. It is hoped that this study will be able to reproduce previously seen spectral characteristics using FTIR spectroscopy and use that information to extract a more definite result using Raman spectroscopy. However, before dealing with the detailed interpretation of tissue spectra with a view to highlighting cancer, the following section will investigate the spectra of the constituents of the biological macromolecules that go on to form human cells and tissues. This will provide a better understanding of the spectra of cells and tissues.

Chapter 2

Raman Spectroscopy of Biological Molecules and Tissue

2.1 Introduction

As discussed in Chapter 1, biological macromolecules and cellular components are made up of smaller molecules, such as amino acids, that bond together in successively elaborate steps to form the intricate molecules seen in biological systems (Fig. 2.1). In order to examine biological systems, Raman spectroscopy was chosen for this work due to the high spectral and spatial resolution it afforded, as well as the minimal sample preparation it required. Before examining tissue, Raman spectroscopy was first employed to examine the component molecules of tissue and the effects of increasing molecular complexity on the resulting Raman spectrum. In the detection of biochemical changes in cancerous tissue any changes in the overall tissue may be attributable to changes in one or any of its components. This study aims to investigate the quality of Raman spectra obtained from biological material and the degree to which component molecules can be resolved from a heterogeneous sample. This was achieved by recording Raman spectra from amino acids, peptides, proteins, purines, pyrimidines, DNA, RNA, larger biological macromolecules and finally tissues

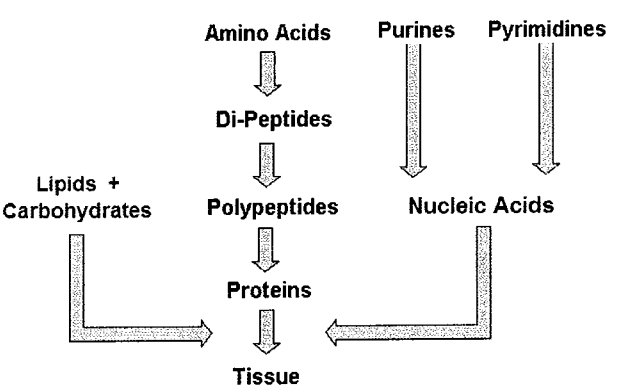


Figure 2.1 Increasing levels of biological complexity

2.2 Materials and methods

2.2.1 Instrumentation

The Instruments S.A. Labram Raman spectroscopic confocal microscope was used, with an Argon Ion laser operating at an excitation wavelength of 514.5nm, as discussed in section 1.3.4. A x50 objective lens was used. The laser power at the sample was measured and found to be 6.5±0.05mW, focused to a spot size of ≈2μm at the tissue surface.

2.2.2 Sample preparation

The samples examined were:

1. 20 amino acids
2. 2 purines; Adenine and Guanine
3. 2 pyrimidines; Cytosine and Thymine
4. Di-peptide; Arginine-Lysine
5. Polypeptide; Kemptide (Leu-Arg-Arg-Ala-Ser-Leu-Gly)
6. Proteins; Trypsin, Albumin and β-Galactose
7. Salmon DNA and Bovine RNA
8. Glucose, Glycogen and Collagen (calf skin)

9. Tissues; Formalin fixed paraffin preserved (FFPP) sections including; Hodgkins lymphoma, tonsil, naevi, melanoma, thyroid, medullary thyroid cancer, kidney and renal tumour

Spectra of samples 1-9 (Sigma-Aldrich, Dublin, Ireland) were recorded in a polycrystalline form on a glass slide with no pre-treatment. Due to the purity of samples 1-8, a 30-second accumulation proved sufficient.

FFPP tissue samples were obtained from both the School of Biological Sciences, DIT, Kevin Street and the National Maternity Hospital, Holles St., Dublin. FFPP sections were sliced into 5μm sections using a microtome, mounted on glass cover slips and dried. Samples were dewaxed in-house in DIT prior to investigation by immersion in baths of Xylene (BDH), Ethanol Absolut (Merck) and Industrial Methylated Spirits 95% (Lennox) and air-dried. Raman spectra of the various tissues were accumulated for 150 seconds.

Selected spectra were baseline corrected, dark current subtracted and where multiple spectra are displayed, spectra have been offset to facilitate clarity.

2.3 Results

2.3.1 Amino acids, peptides and proteins

The Raman spectra of the amino acids show many narrow bands, as would be expected due to their relatively simple structure. Each of the Raman peaks corresponds to a vibrational energy associated with the various bonds within the sample. The spectra display the ability of Raman spectroscopy to classify each of these molecules with their own distinctive signature spectrum. Although all 20 common amino acids were examined only a selection are shown here. However all of the amino acids produced their own distinct good quality spectra. Figure 2.2 shows the chemical structure and the Raman spectrum of phenylalanine. A selection of peaks corresponding to vibrations of bonds within the molecule have been highlighted. The most distinctive peak in the spectrum of phenylalanine is the peak at 1004 cm^{-1} which can be attributed to the carbon-carbon stretching mode within the phenyl ring. Peak assignments in the following sections were made using Socrates (2004), Lakshmi *et al.* (2002) and Barry *et al.* (1992). The spectra of amino acids, purines and pyrimidines correspond closely with previously recorded spectra available in the Spectral Database of Organic Compounds (2006).

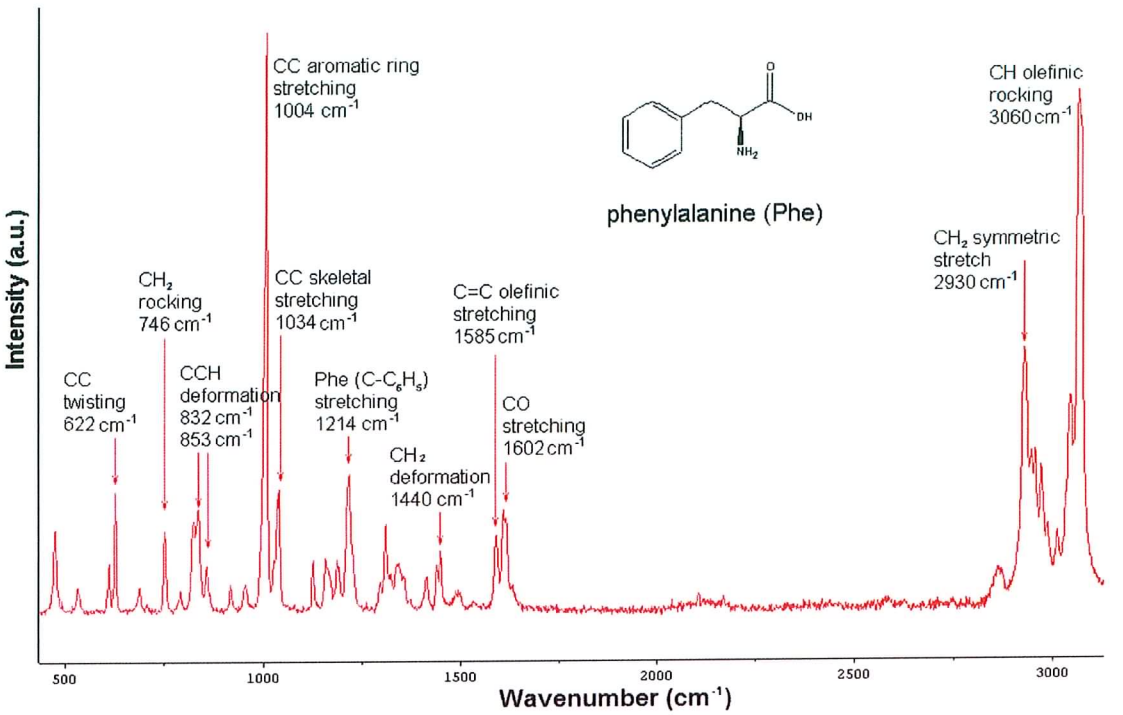


Figure 2.2 Raman spectrum of phenylalanine powder (hydrophobic amino acid)

The contribution at 1214 cm^{-1} is a peak resulting from stretching between the backbone carbon and the phenyl ring (C-C_6H_5). Other peaks are assigned to C-C twisting (622 cm^{-1}), C-C stretching (1034 cm^{-1}), C=C stretching (1585 cm^{-1}), CH_2 rocking (746 cm^{-1}), CH_2 scissoring (1440 cm^{-1}), CH_2 stretching (2930 cm^{-1}) and CO stretching (1602 cm^{-1}) modes. Figure 2.2 shows how the fingerprint region of the spectrum (≈ 500 - 1700 cm^{-1}) contains the majority of information about this sample and above 1700 cm^{-1} the spectrum contains mainly CH, CH_2 and CH_3 stretching modes around the 3000 cm^{-1} region. In order to reduce the acquisition time and minimise the tissue's exposure to potential damage resulting from the laser, the fingerprint region alone was recorded throughout the present study.

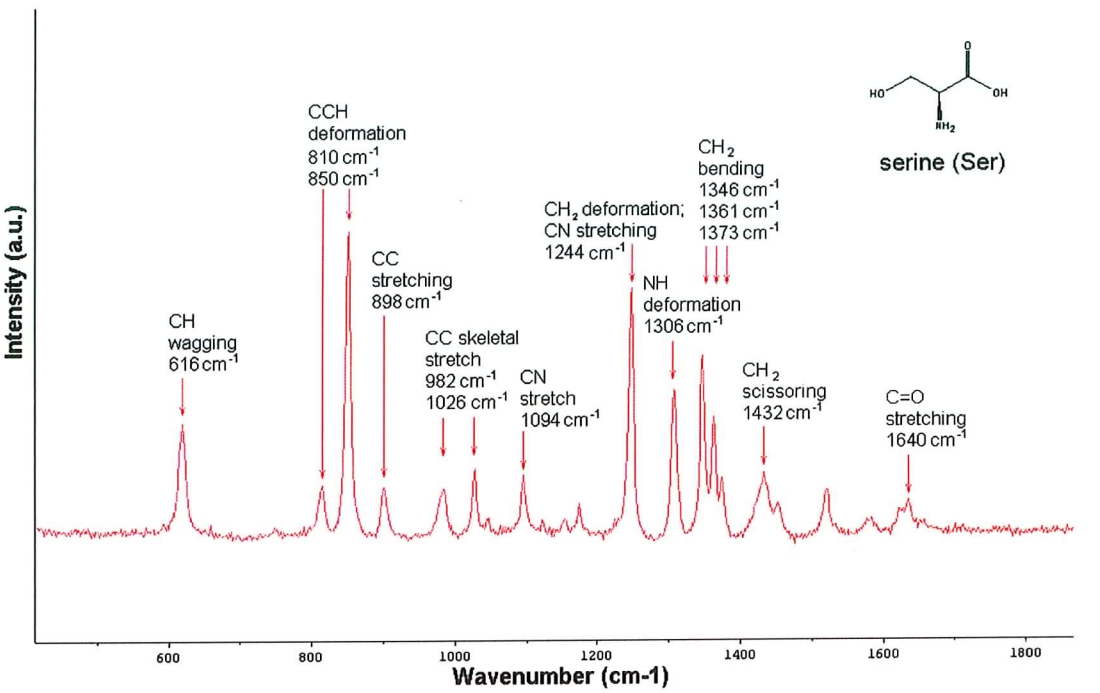


Figure 2.3 Raman spectrum of serine (hydrophilic amino acid)

Figure 2.3 shows the structure and Raman spectrum of serine (a member of the hydrophilic group of amino acids). Contributions are observed from CCH deformation (810 cm^{-1} and 850 cm^{-1}), CH wagging (616 cm^{-1}), CH_2 wagging, bending and scissoring (1244 cm^{-1}, 1346-1373 cm^{-1} and 1432 cm^{-1} respectively), CC stretching (898 cm^{-1}), CC skeletal stretching (982 cm^{-1} and 1026 cm^{-1}), C=O stretching (1640 cm^{-1}), CN stretching (1094 $cm^{-1}$and 1244 cm^{-1}) and NH deformation (1306 cm^{-1}) modes.

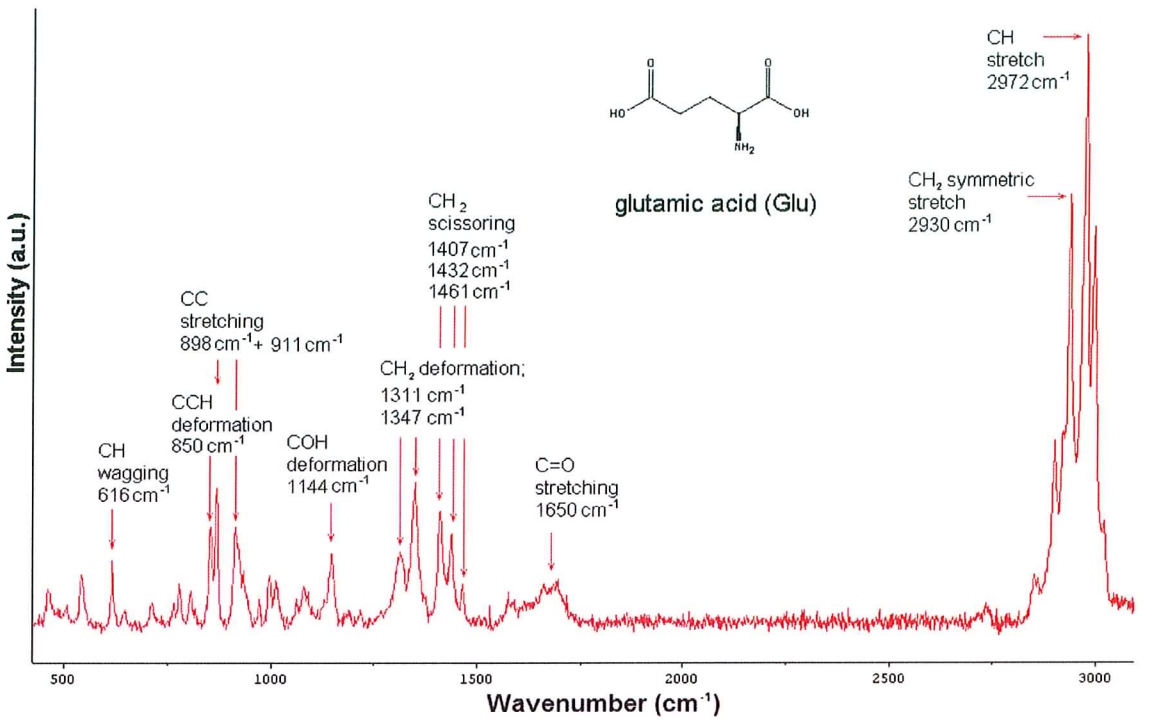


Figure 2.4 Raman spectrum of glutamic acid (acidic/ negatively charged amino acid)

The Raman spectrum of glutamic acid (Fig. 2.4) exhibited Raman contributions from CH wagging (616 cm^{-1}), CCH deformation (850 cm^{-1}), CC stretching (898 cm^{-1} and 911 cm^{-1}), COH deformation (1144 cm^{-1}), CH_2 twisting and scissoring (1311 - 1461 cm^{-1}), C=O stretching (1650 cm^{-1}), CH_2 symmetric stretch (2930 cm^{-1}) and CH stretching (2972 cm^{-1}) modes.

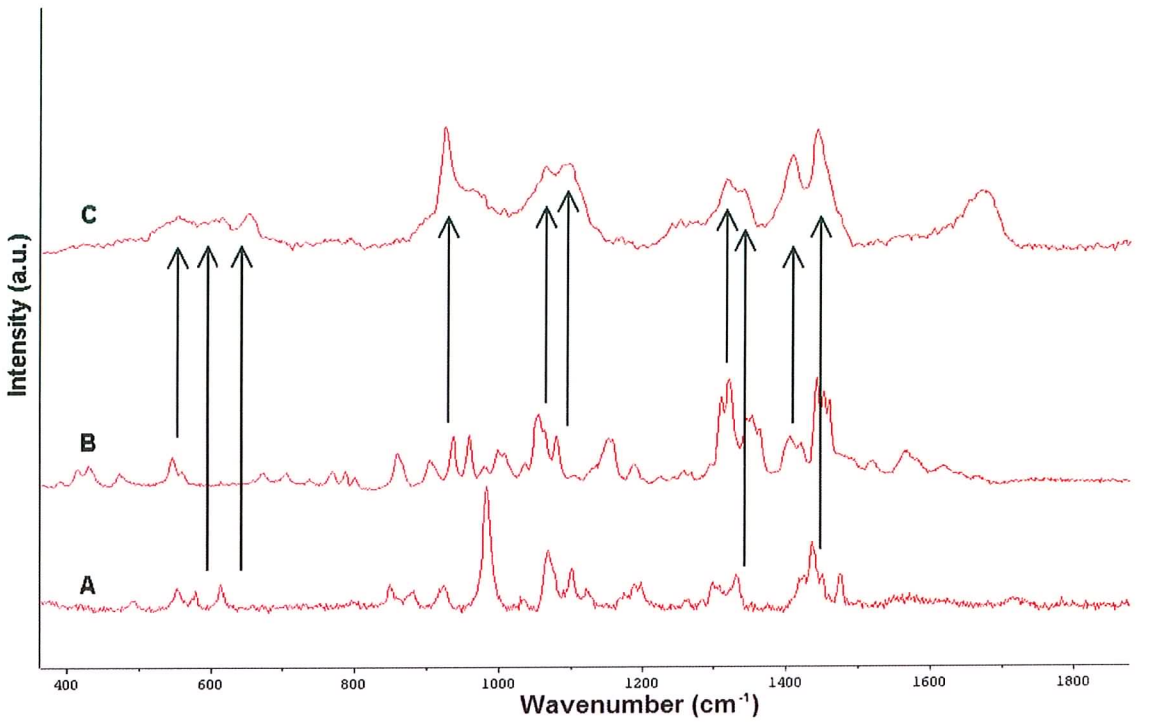


Figure 2.5 Raman spectra of A) arginine B) lysine and C) dipeptide formed between arginine and lysine

Figure 2.5 compares the Raman spectra of arginine (Arg) and lysine (Lys) to the Raman spectrum of the dipeptide Arg-Lys. The peptide Arg-Lys is simply the two amino acids connected by a peptide bond. The spectrum clearly derives from combinations of the two amino acids (indicated by the arrows), but is more complex than a simple sum of the parts. Figure 2.6 displays the spectrum of the dipeptide with assignments.

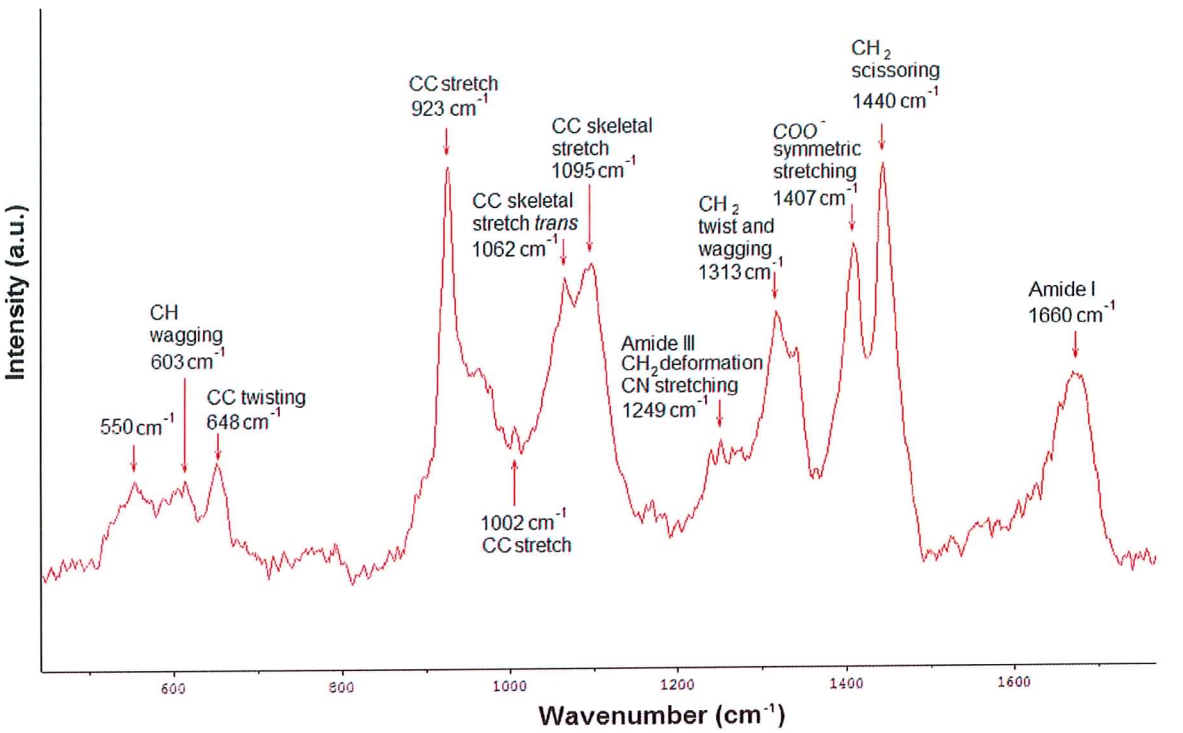


Figure 2.6 Raman spectrum of Arg-Lys with peak assignments

The vibrational modes seen in the dipeptide include CC stretch, CH_2 deformation, amide I and III bands. Amide I bands are mainly due to C=O stretching modes of the peptide bond and amide III bands are mainly due to in-plane NH bending and CN stretching modes. It was found that in this very simple first step in the formation of a dipeptide its spectrum displayed many of the same features as seen in spectra of more complex molecules. Figure 2.7 shows the Raman spectrum of the polypeptide kemptide. Although kemptide contains five amino acids (leucine (Leu), arginine, alanine (Ala), serine (Ser) and glycine (Gly)) its Raman spectrum is remarkably similar to that of Arg-Lys. However at this point is becomes difficult to assign peaks resulting from a single component, as each amino acid in the chain has many of the same bonds and as a result the peaks become broader. The spectrum at this

point represents a snapshot of the ensemble of constituent moieties contained within the polypeptide.

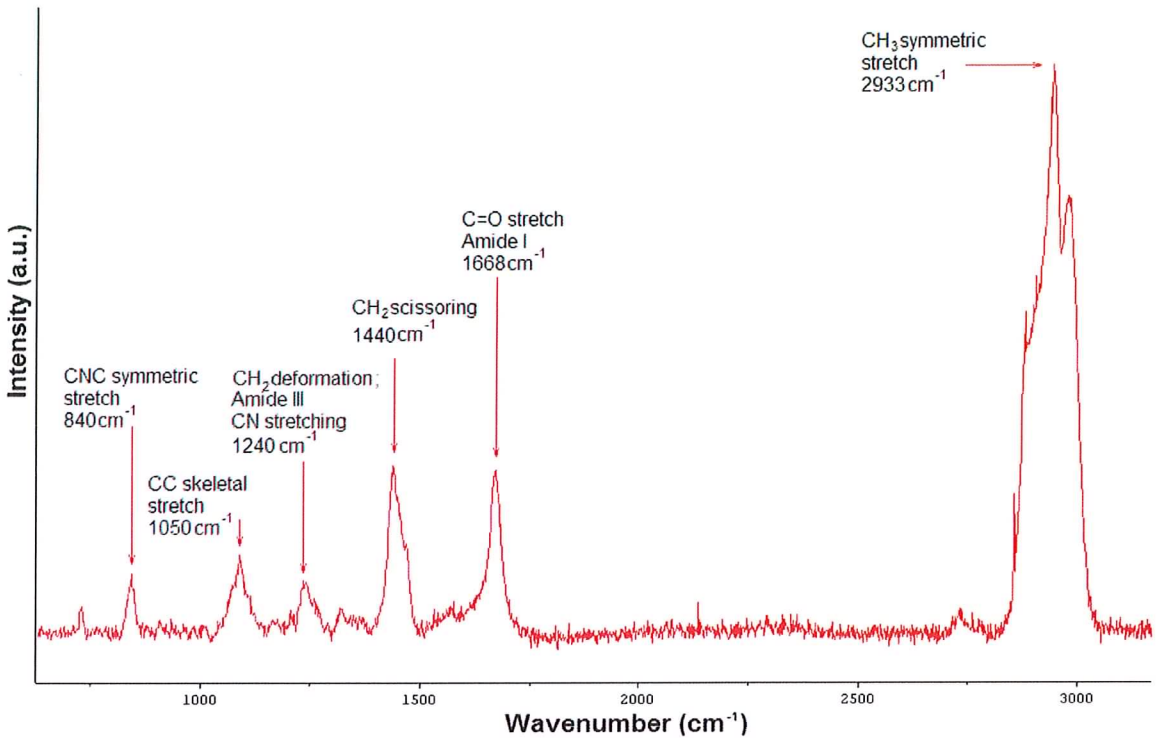


Figure 2.7 Raman spectrum of the polypeptide kemptide (Leu-Arg-Arg-Ala-Ser-Leu-Gly)

It is sequences of polypeptides, such as in kemptide, which combine into long strands to form proteins. Three different proteins were examined with increasing molecular weights and hence, chain lengths. The proteins examined (Figure 2.8) were Trypsin, Albumin and β-Galactose, with molecular weights of 20,000, 65,000 and 115,000 Dalton respectively (unit of weight is the Dalton, which is one-twelfth the weight of an atom of ^{12}C). Figure 2.8 demonstrates the similarities in the spectra between polypeptides and proteins. When the amino acids combine to make a peptide a water molecule is lost and the peptide bond is formed. These peptide bonds are very strong bonds and proteins are merely increasingly

long chains of peptides. The morphology of the proteins is determined by the secondary structure. Changes in the secondary structure of macromolecules results in shifts and changes in the shapes of the Raman bands, although specific contributions from individual vibrations (i.e. CC aromatic ring breathing, 1004 cm^{-1}) are still present. β-Galactose is the longest of the three proteins studied and clearly its spectrum has more components than the spectra of the shorter protein chains.

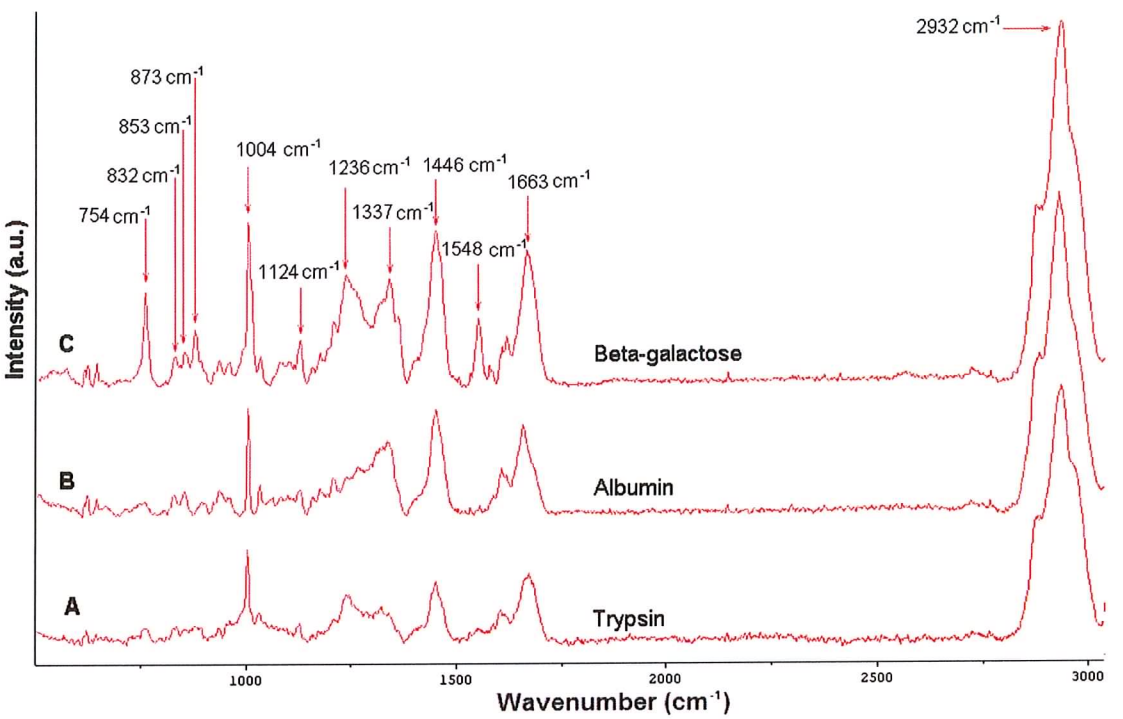


Figure 2.8 Raman spectra of three proteins in ascending chain length; A) Trypsin, B) Albumin and C) β-Galactose.

Peaks identified include;

Symmetric ring breathing (754 cm^{-1}),

CCH deformation aliphatic and aromatic (832 cm^{-1} and 853 cm^{-1} respectively),

CC stretch (873 cm^{-1}),

CC aromatic ring breathing mode (1004 cm^{-1}),

CC skeletal stretch trans conformation (1124 cm^{-1}),

CN stretch, NH bending Amide III band (1236 cm^{-1}),

CH_2 deformation (1337 cm^{-1}),

CH_2 scissoring (1446 cm^{-1}),

NH deformation; CN stretch Amide II band (1548 cm^{-1}),

C=O stretch Amide I α-helix (1663 cm^{-1}),

CH_3 symmetric stretch (2932 cm^{-1}).

2.3.2 Purines, pyrimidines, DNA and RNA

A similar investigation was carried out into a selection of purines and pyrimidines. Figure 2.9 displays the chemical structure of DNA. The Raman spectra of the purines and pyrimidines present in DNA, as well as a spectrum taken from salmon DNA are shown in Figure 2.10. Again contributions from the constituent molecules are present in the spectrum of DNA, but as seen in proteins the peaks broaden as the molecular complexity increases. The spectrum of DNA corresponds with previously published results (Kneipp *et al.* 2003).

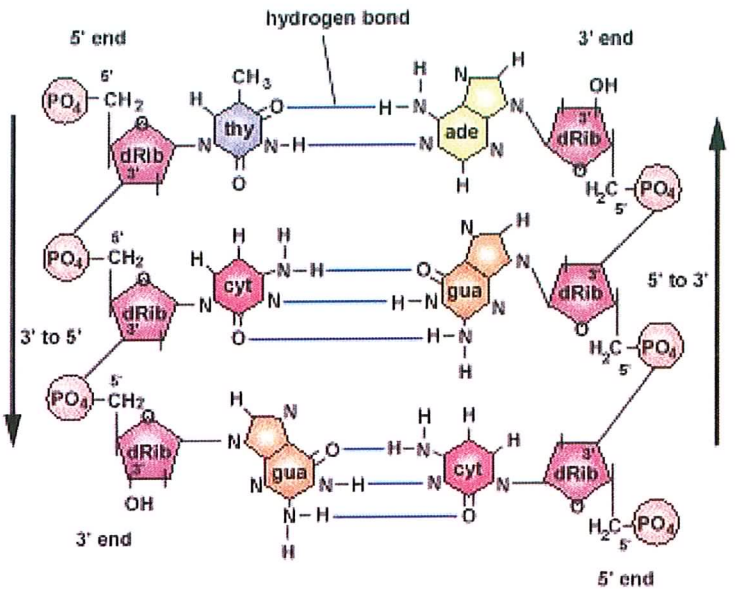


Figure 2.9 Chemical structure of DNA

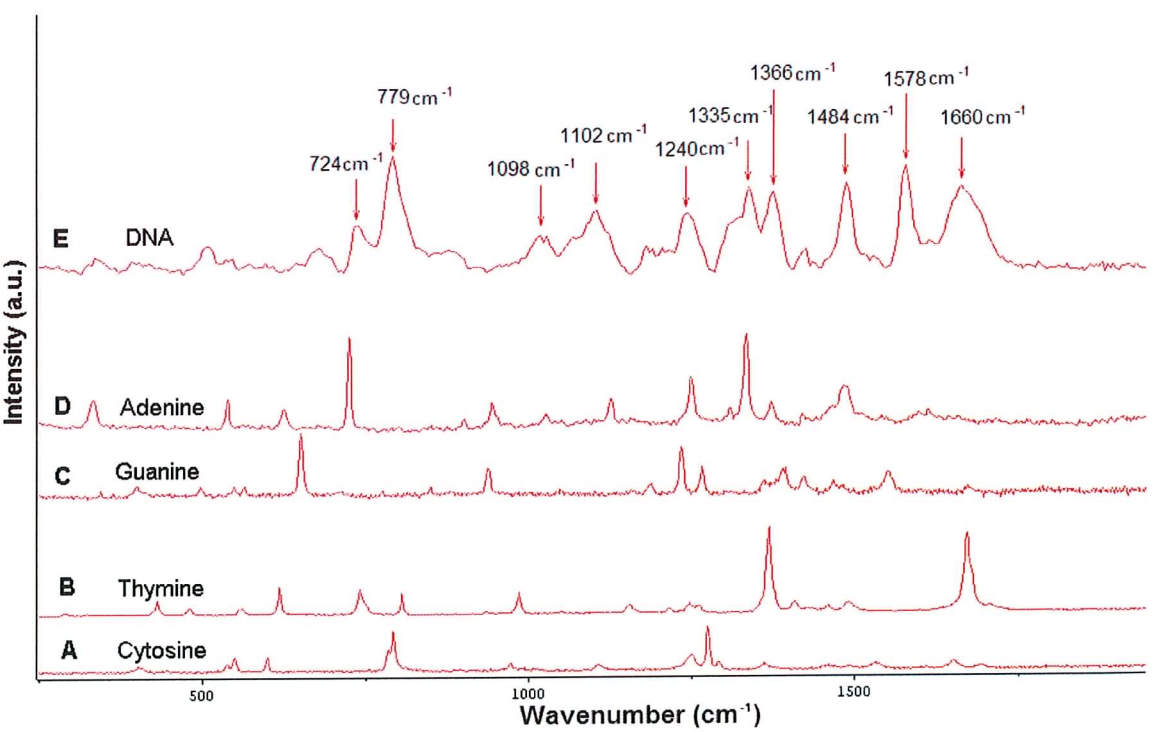


Figure 2.10 Raman spectra of A) Cytosine B) Thymine (pyrimidines),

C) Adenine D) Guanine (purines) and E) Salmon DNA

Peaks identified in DNA include;

CH_2 deformation (724 cm^{-1}),

Ring vibration (779 cm^{-1}),

CC stretch (1098 cm^{-1}),

CC stretch (1102 cm^{-1}),

CN stretch, NH bending Amide III band (1240 cm^{-1}),

CH_2 deformation (1335 cm^{-1}),

CH_2 bending (1366 cm^{-1}),

CH_2 deformation (1484 cm^{-1}),

C=C olefinic stretch (1578 cm^{-1}),

C=O stretch Amide I α-helix (1660 cm^{-1}).

2.3.3 Biological Macromolecules

Raman spectra were also recorded from biological macromolecules (DNA, RNA, glucose, glycogen and collagen) and results compare well with previously published data (de Jong *et al.* 2003, Kneipp *et al.* 2003, de Jong *et al.* 2002, Koljenović *et al.* 2002). The spectra of DNA and RNA (Figure 2.11) show remarkable similarity, which is to be expected given their similar molecular structure. However a peak at 809 cm^{-1} appears more prominently in DNA than in RNA, this peak may have the potential to differentiate between RNA and DNA.

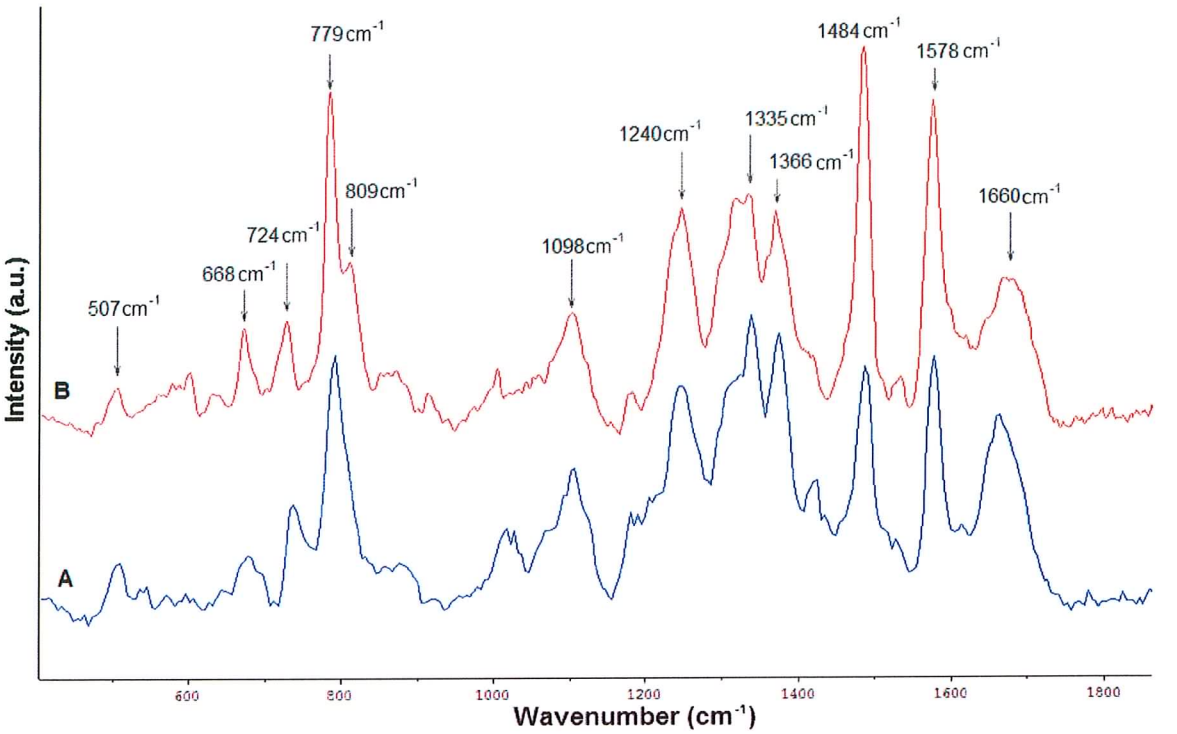


Figure 2.11 Raman spectra of A) DNA and B) RNA

Peaks observed include;

507 cm^{-1} unassigned

668 cm^{-1} unassigned

CH_2 deformation (724 cm^{-1}),

O-P-O stretch (788 cm^{-1}),

PO_2^- nucleic acid backbone (1098 cm^{-1}),

CN stretch, NH bending Amide III band (1240 cm^{-1}),

CH_2 deformation (1335 cm^{-1}),

CH_2 bending (1366 cm^{-1}),

CH_2 deformation (1484 cm^{-1}),

C=C olefinic stretch (1578 cm^{-1}),

C=O stretch Amide I α-helix (1660 cm^{-1}).

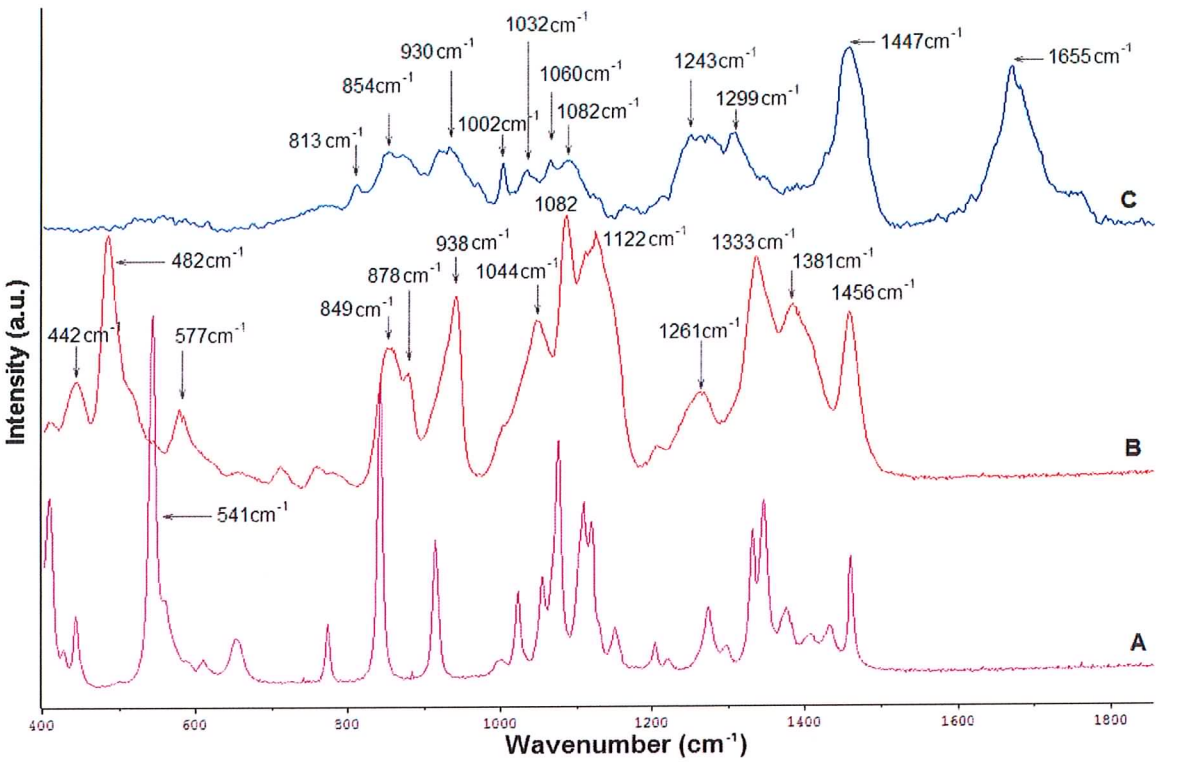


Figure 2.12 Raman spectra of A) Glucose, B) Glycogen and C) Collagen

Figure 2.12 shows the Raman spectra of glucose, glycogen and collagen. Glucose displays many narrow vibrational modes, similar to the spectra of single amino acids, whereas the spectrum of glycogen shows broader spectroscopic features.

Peaks identified in glycogen include;

CC skeletal deformation (442 cm^{-1}),

CC skeletal deformation (482 cm^{-1}),

577 cm^{-1} unassigned

CCH aromatic deformation (849 cm^{-1}),

CH_2 deformation (878 cm^{-1}),

CCH deformation (938 cm^{-1}),

C-O stretch (1044 cm^{-1}),

C-C stretch (1082 cm^{-1}),

C-C stretch (1122 cm^{-1}),

CH_2 wagging (1261 cm^{-1}),

CH_3CH_2 wagging (1333 cm^{-1}),

OH deformation (1381 cm^{-1}),

CH deformation (1456 cm^{-1}),

Peaks in collagen included;

CCH aliphatic deformation (813 cm^{-1}),

CCH aromatic deformation (854 cm^{-1}),

CC stretch α-helix (930 cm^{-1}),

CC aromatic ring stretch (1002 cm^{-1}),

CC skeletal stretch (1032 cm^{-1}),

CC skeletal stretch, *trans* conformation (1060 cm^{-1}),

CC skeletal stretch, random conformation (1082 cm^{-1}),

CH_2 wagging (1243 cm^{-1}),

CH_2 deformation (1299 cm^{-1}),

CH_2 scissoring (1447 cm^{-1}),

C=O stretch Amide I (1655 cm^{-1}),

With spectra recorded from a selection of biological macromolecules found in human tissue, next spectrometer setting were optimised for recording spectra from human tissue sections.

2.3.4 Spectrometer Settings for Recording Tissue

Investigations were then carried out into the components of cells and tissue. Tissue was obtained from dewaxed formalin fixed paraffin preserved (FFPP) sections. However unlike powders examined previously, tissue needed to be handled with particular care due to the possibility of damage from the laser light. The Raman spectrometer has a number of variables that must be optimised in order to give the best signal possible, while ensuring that no damage is caused to the sample by laser heating effects. These variables are; slit size, hole size, laser power, spot size, grating and spectral range. A change in one of these variables can have an effect on the sample being examined, as well as the quality of the spectrum obtained. Many different settings and combinations of setting were tried. However the settings chosen as shown in Table 2.1 resulted from a compromise between the best Raman spectra attainable and minimum exposure / damage to the samples under investigation.

Hole Size	600 μm
Slit Size	920 μm
Laser Power	6.5±0.05mW
Spot Size	≈2μm

Table 2.1 Final optimised Raman settings used for tissue samples

Once optimisation of the experimental setup had been achieved and acceptable Raman spectra were being obtained from tissue samples, a suitable protocol for recording spectra was required. Although the 600 lines/mm grating was initially used, the 1800 lines/mm grating improved the resolution from 3.3cm^{-1} per pixel to 1.65cm^{-1} per pixel (Fig. 2.13).

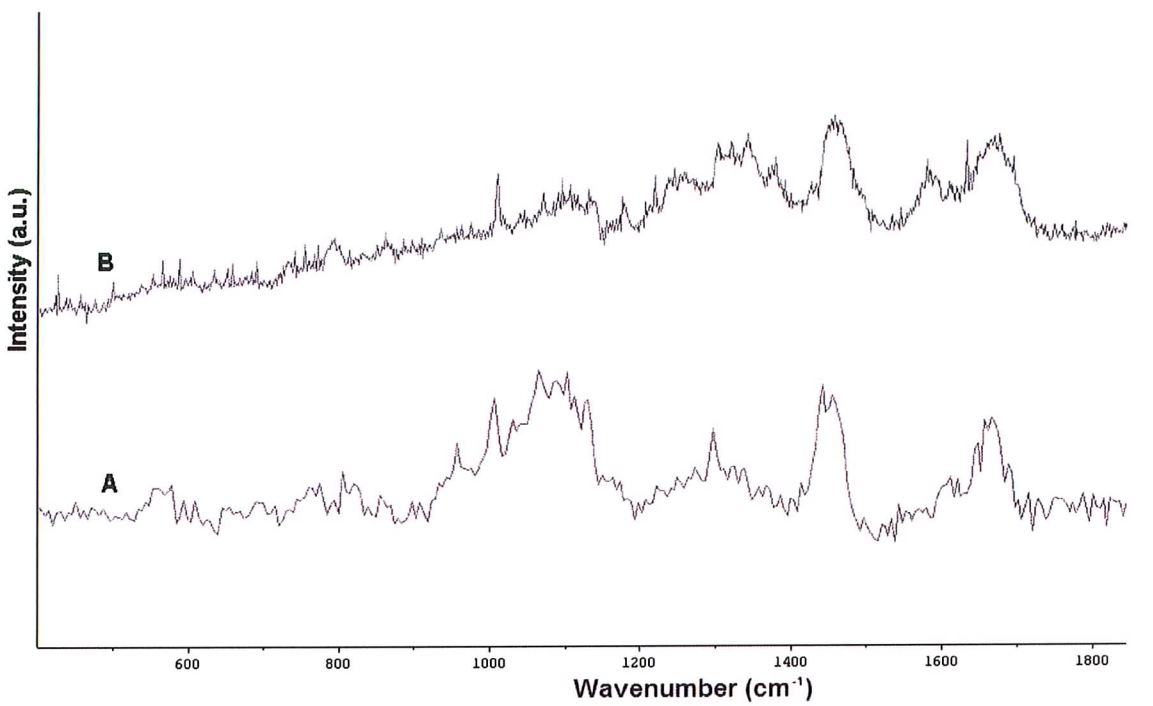


Figure 2.13 Raman spectra of an FFPP tissue section obtained using (A) 600 lines/mm (B) 1800 lines/mm gratings

Silicon was used to calibrate the spectrometer by setting its signature peak (Fig. 1.17) to 520.7 cm^{-1}. Raman spectra of the tissue section were measured from 400 to 1900 cm^{-1} relative to the excitation frequency. This is the spectral region most frequently examined in literature with respect to tissue specimens. Various collection times were also tested and an accumulation time of 150 seconds was found to be optimal, although spectra were also

taken for 10 and 60 seconds (Fig. 2.14). These multiple measurements allowed the spectra to be monitored for spectral changes as the accumulation time progressively increased.

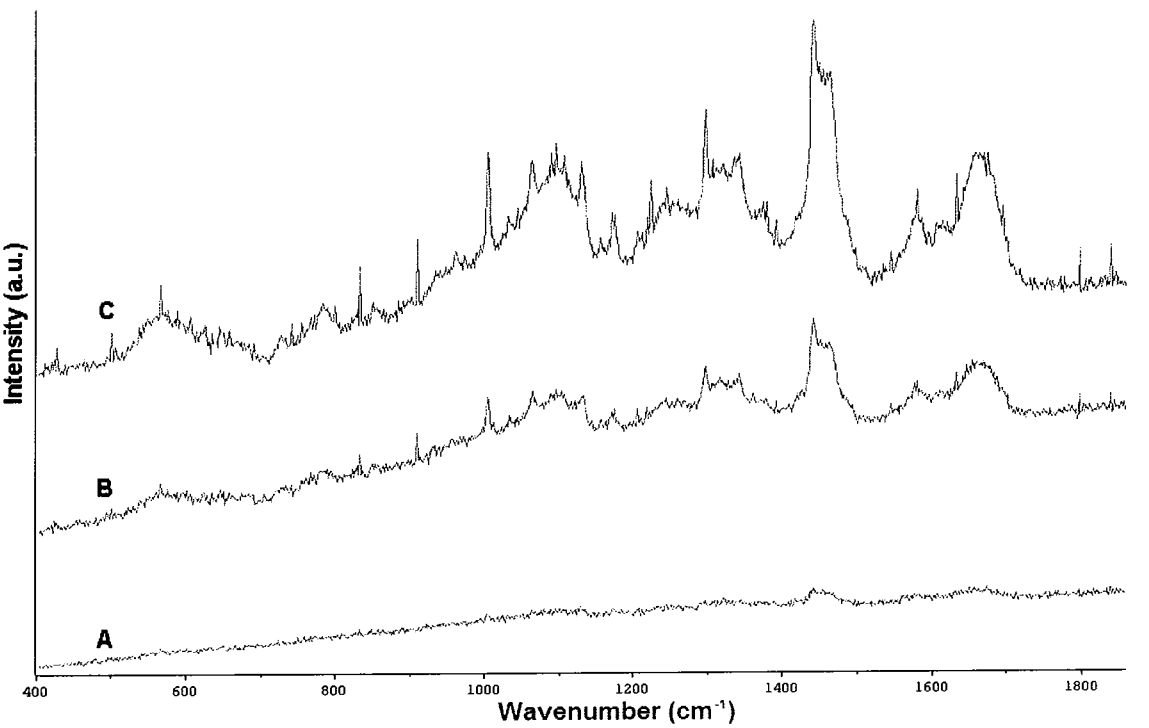


Figure 2.14 Raman spectra of an FFPP tissue section at three accumulation times A)10 sec B) 60 sec and C) 150sec, using 1800 lines/mm grating and spectral range of 400–1900 cm^{-1}

The effect of using the various objective lenses was also investigated. Using the X10 and X20 objective lenses proved unsuccessful due to their low numerical aperture, as insufficient Raman signal was being collected. However the X50 and X100 objectives both produced similar spectra (Fig. 2.15), but the X100 sometimes resulted in sample burning over the 150sec acquisition and as a result the X50 lens was used. Using the X50 lens, the laser power at the sample was measured and found to be 6.5±0.05mW, focused to a spot size of ≈2μm at the tissue surface.

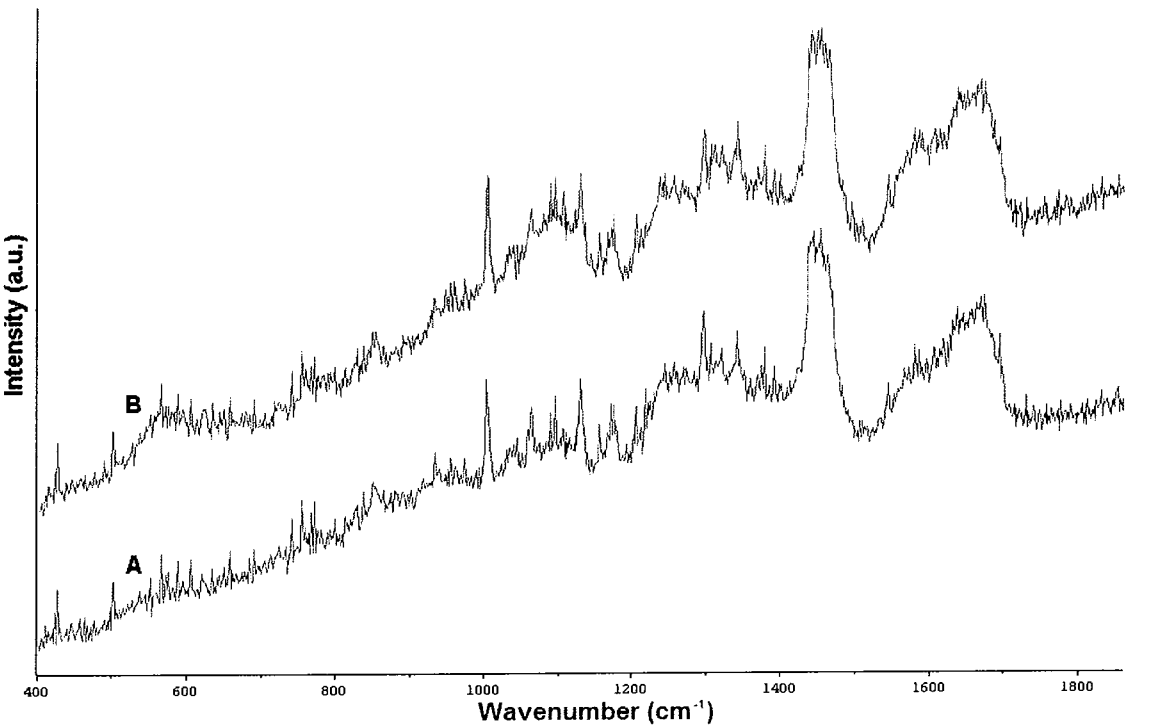


Figure 2.15 Comparison of Raman spectra of an FFPP tissue section with A) X100 lens and B) X50 lens

Figure 2.15 also displays the commonly observed background encountered with 514.5 nm laser light. This background may have contributions from fluorescence and/or stray scattered light. Although the background is evident in the spectra, the Raman signal is clearly discernible above it and simple baseline subtraction produces good quality spectra.

2.3.5 Tissue

A large variety of tissues were examined. A selection is shown in Figure 2.16. It is clear that there are some vibrational regions common in much of the samples examined. The

Amide I band evident in proteins contributes strongly as do CH_2 bending, CH_2 deformation, Amide III vibrations, CC stretching and ring breathing modes. These have been outlined and assigned in Table 2.2.

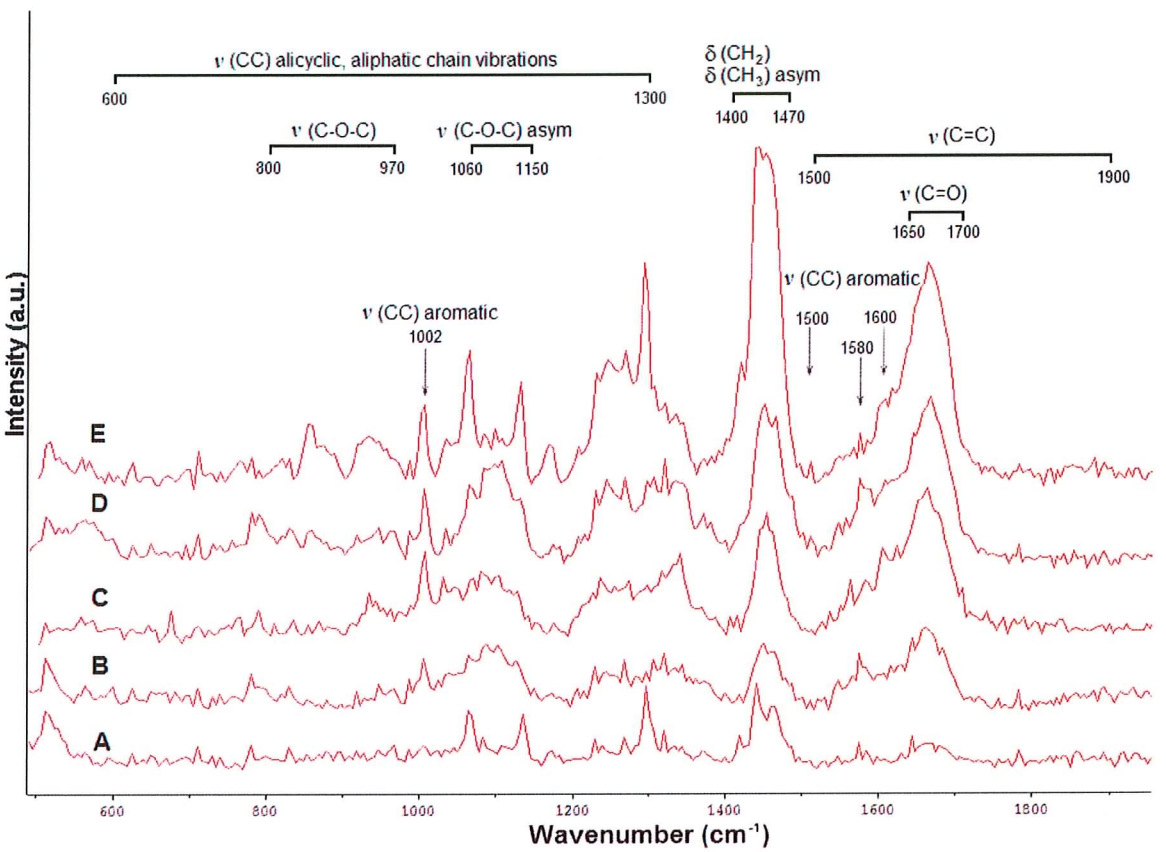


Figure 2.16 Common spectral features of various tissues, A) Tonsil B) Hodgkin's lymphoma and C) Renal tumour D) Kidney and E) Prostate,

where ν = stretch and δ = deformation

Spectral region	General Assignment
≈1000 cm^{-1}	Ring breathing
≈1000 – 1200 cm^{-1}	CC stretching
≈1200 – 1500 cm^{-1}	CH_2 deformation and Amide III band
≈1500 – 1600 cm^{-1}	C=C stretching
≈1600 – 1700 cm^{-1}	Amide I band

Table 2.2 Outline of general Raman spectral assignments for tissue samples

Figure 2.17 compares the Raman spectrum of cervical tissue with the Raman spectrum of glass. Due to the fact that tissue sections are very thin (5μm), spectral contributions were observed from the underlying glass slide which gave two broad peaks at 560cm^{-1} and 1086cm^{-1}. However Raman peaks from tissue in these areas are still clearly visible on top of these signals.

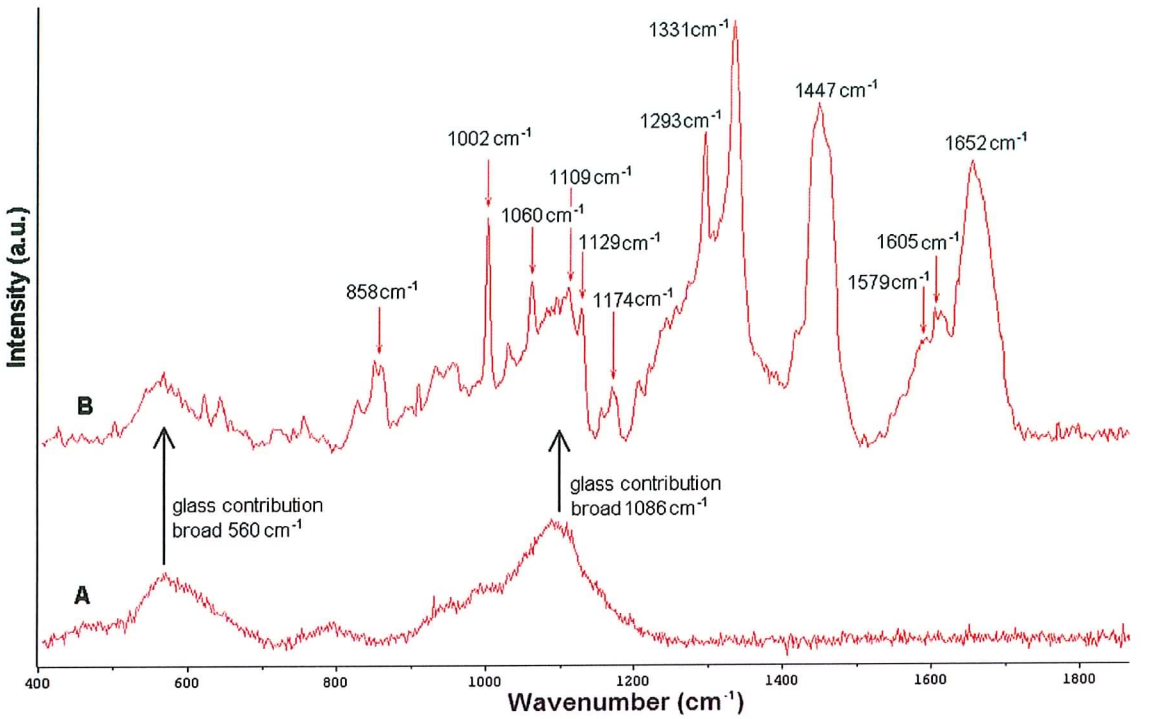


Figure 2.17 Raman spectra of A) a glass slide and B) dewaxed FFPP cervical section on glass

The main spectral contributions with assignments are summarized as follows;

CCH deformation aliphatic and aromatic (832 cm^{-1} and 853 cm^{-1} respectively),

CC aromatic ring breathing mode (1002 cm^{-1}),

CC skeletal trans-conformation (1060 cm^{-1}),

CC stretching (1109 cm^{-1}),

CC skeletal stretch (1129 cm^{-1}),

CC stretching (1174 cm^{-1}),

CH_2 deformation (1293 cm^{-1}),

CH_2 deformation (1331 cm^{-1}),

CH_2 deformation (1447 cm^{-1}),

C=C olefinic stretch (1579 cm^{-1}),

C=C in-plane bending (1605 cm^{-1}),

C=O stretch Amide I α-helix (1652 cm^{-1}).

2.4 Conclusion

Raman spectra were successfully obtained from all the specimens examined. The smallest molecules displayed much fine spectral detail due to their relatively simple chemical structure. The bonds in these amino acids could be identified from their Raman spectra. Peptides also yielded good Raman spectra and the details of the constituent amino acids could be resolved from the overall spectrum.

It is clear that many of the characteristic bands of amino acids are also present in the spectra of dipeptides, although much of the narrow bands and fine structure associated with the simple amino acid has been replaced by a broader structure in the peptide. On closer examination of the spectrum of the peptide Arg-Lys, the individual peaks can be assigned to molecular vibrations of constituent bonds. The Amide I band appears as a result of the peptide bond after the addition of only two amino acids.

On examination of a polypeptide (Kemptide) a reduction in the fine spectral structure was observed. The polypeptide showed only five or six broad contributions. Like the polypeptide the three proteins examined (Trypsin, Albumin and β-Galactose) displayed a small number of broad peaks, increasing in complexity and intensity with the increasing chain length of the protein.

The constituent purines and pyrimidines of DNA were also examined (cytosine, thymine, adenine and guanine) and as with the peptides and polypeptides, the constituent components could be resolved from a Raman spectrum of pure DNA. Biological macromolecules such as DNA, RNA, glycogen and collagen were examined and each produced good quality and distinct signature spectrum. Conditions were optimised to ensure there was no sample deterioration in the tissue sections. A x50 objective lens, resulting in 6.5±0.05mW at the sample surface for 150 seconds accumulations was employed.

Having optimised the Raman parameters to suit the tissue specimens, good quality Raman spectra, comparable with published data were obtained. Although the 514.5nm laser line did produce a degree of background, it did not obscure spectral features of biological material. The parameters outlined in this section have been used throughout this study.

Raman spectra of tissue had peaks similar to that of proteins, however the overall spectrum was more complicated than protein alone. A large variety of tissue was examined, including; Hodgkins lymphoma, tonsil, naevi, melanoma, thyroid, medullary thyroid cancer, kidney, renal tumour. A number of main vibrational regions were identified, namely, ring breathing, CC stretching, CH_2 deformation, Amide III band, C=C stretching and Amide I band. These bands were consistently found in the different tissue types examined.

Glass contributions were seen in the spectra recorded from 5μm sections. To minimise this unwanted spectral feature, 10μm sections were obtained for subsequent studies.

Chapter 3

Evaluation of the Effect of Tissue Processing

3.1 Introduction

In the past number of years many studies have been carried out using vibrational spectroscopy to classify tissue with a view to cancer diagnosis. Both IR and Raman spectroscopies have been employed. As outlined in chapter 1, Raman spectroscopy is very similar to the more frequently used Fourier Transform Infrared (FTIR) spectroscopic technique and both techniques are very complementary. The work presented in this chapter has been published in the Journal of Vibrational Spectroscopy (Ó Faoláin *et al.* 2005a).

Some of the tissue types examined by various groups include; **cervical** (Utzinger *et al.* 2001, Mahadevan-Jansen *et al.* 1998, Yazdi *et al.* 1998, Liu *et al.* 1992), **breast** (Smith *et al.* 2003, Kneipp *et al.* 2003, Shah *et al.* 2001,), **skin** (McIntosh *et al.* 2002,Nunes *et al.* 2003, Nijssen *et al.* 2002, Caspers *et al.* 2001, Gniadecka *et al.* 1997), **lung** (Kaminaka *et al.* 2001), **brain** (Mizuno *et al.* 1994, Keller *et al.* 1994), **bladder** (Crow *et al.* 2003), **eosophagus** (Boere *et al.* 2003, Kendall *et al.* 2003, Song *et al.* 2003), **colon** (Molckovsky *et al.* 2003), **liver** (Chiriboga *et al.* 2000, Hawi *et al.* 1996), **thyroid** (Liu *et al.* 2003) and **prostate** (Crow *et al.* 2003, Paluszkiewicz *et al.* 2001) . A variety of different methods of sample preparation have been employed in these studies, such as; fresh, frozen, air dried, formalin fixed and dewaxed FFPP tissue sections. The majority of tissue studies have been carried out on either frozen tissue or dewaxed FFPP sections (Dukor 2002). Although considered the gold-standard, frozen tissue is difficult to obtain, deteriorates rapidly and cannot provide retrospective analysis. However there is a plentiful supply of archival tissue

samples, preserved and mounted in wax blocks. Due to the prevalence of these FFPP tissue sections, a better understanding of the effects of processing could unlock the potential diagnostic capabilities of these sections. The sections go through many processing steps before spectroscopic evaluation. If these processing steps result in changes in spectral bands identified as diagnostically significant, this would have implications for tissue classification and could compromise the potential of these FFPP sections in a diagnostic capacity.

Previous studies using vibrational spectroscopy have been carried out into the effect of formalin fixation as well as ex-vivo handling (Huang *et al.* 2003, Shim *et al.* 1996, Pleshko *et al.* 1992). Huang *et al.* (2003) investigated the effects of formalin fixation on human bronchial tissue using Raman spectroscopy. A decrease in overall Raman intensities was observed and notable formalin peaks were identified at 907, 1041 and 1492 cm^{-1}. Shim *et al.* (1996) examined the effect of tissue drying, formalin fixation and snap freezing on normal hamster tissues using Raman spectroscopy. The study found that formalin did not contaminate the Raman spectrum (with the exception of the 1041 cm^{-1} band), however drying/dehydration was found to disrupt protein vibrational modes. FTIR spectroscopy was used by Pleshko *et al.* (1992) to study the effects of fixation in ethanol and formalin, as well as methacrylate embedding. The study found ethanol fixation of fetal rat bone tissue resulted in changes in the Amide I and II bands (1650 and 1550 cm^{-1} respectively), a result of alteration of the protein conformation of the tissue.

The aim of this study is to re-examine the above and extend the study to investigate the effect of xylene, paraffin wax embedding and subsequent dewaxing on human tissue using

both Raman and FTIR spectroscopies. The use of both techniques on the same samples maximises the amount of structural information obtained from the tissue. The steps involved in processing these sections can be summarised as follows: tissue excising, fixation in formalin, tissue dehydration in alcohol, embedding in paraffin wax, microtomy and removal of wax.

This study examined the biochemical structure of normal parenchymal tissue from the placenta at each of these processing steps using both Raman and IR spectroscopies. This tissue was chosen due to its homogenous nature. This minimises the likelihood of recording spectra from different cell types, which would be expected to produce different spectral features. A homogenous tissue ensures that it is only the effects of processing that are being detected, rather than point to point inhomogeneities.

The effect of tissue freezing was also examined and compared to freshly excised tissue.

3.2 Materials and methods

3.2.1 Raman Instrumentation

The Instruments S.A. Labram Raman spectroscopic confocal microscope was used, with an Argon Ion laser operating at a wavelength of 514.5 nm.. The scattered Raman signal was integrated for 60 to 150 seconds and measured over a spectral range of 400 to 1900 cm^{-1} with respect to the excitation frequency. Once acquired each spectrum was baseline

corrected, dark current subtracted and lightly filtered using a third order linear model to improve clarity.

3.2.2 FTIR Instrumentation

The Perkin Elmer Spectrum GX single-beam, Michelson interferometer based, Fourier Transform Infrared spectrometer was used, which was operated in reflectance mode. The aperture was set to 50 μm × 50 μm, with a gain of 4. Each sample was scanned for 128 scans, baseline corrected and lightly smoothed.

3.2.3 Sample Preparation

Normal parenchymal tissue from the placenta was obtained from the National Maternity Hospital, Holles St., Dublin. Tissue was examined at each stage of the fixation, embedding and dewaxing processes. Tissue was divided into six different pieces and was examined as follows;

1. Fresh tissue: tissue was refrigerated and examined within 24 hrs. The tissue was cross sectioned and a thin section from the tissue centre was excised using a scalpel. The tissue was then rinsed in phosphate buffered saline (PBS), before being mounted onto silver oxide MirrIR slide (MirrIR, Kevley Technologies, Chesterland, Ohio, U.S.) for examination.

2. Frozen tissue: tissue was placed on a metal chuck with optimal cutting temperature (OCT) as support medium and then placed in a cryostat at -20 °C. Tissue was sectioned into 10 μm slices using the cryostat's internal microtome and mounted onto silver oxide MirrIR slides. The frozen sections were placed on the spectrometer stages and allowed to thaw. Tissue spectra were recorded immediately once the tissue thawed.

3. Formalin fixed tissue: tissue was placed in formalin for 24 hrs. The tissue was cross sectioned and a thin section from the tissue centre was excised using a scalpel, rinsed in PBS and mounted onto MirrIR slides before examination.

4. Formalin fixed and dehydrated using xylene (Serosep, Limerick, Ireland): tissue was automatically processed through formalin and xylene. Again the tissue was cross sectioned and a piece of tissue was rinsed in PBS before being mounted on the slides.

5. FFPP tissue mounted in wax: tissue was automatically processed to wax as follows: (i) vacuum fixed in 10% buffered formal saline histo-grade pH 6.8-7.2 (J.T. Baker, Deventer, The Netherlands) and heated to 35 °C (ii) vacuum dehydration in industrial methylated spirit IMS T100 (Lennox, Dublin, Ireland) heated to 35 °C (iii) vacuum clearing in xylene (Serosep, Limerick, Ireland) and heated to 35 °C (iv) vacuum impregnation with Tissue Tek III Embedding Wax with polymer added

(Sakura, Zoeterwoude, The Netherlands) and heated to 59 °C. 10 μm sections were sliced using a microtome and mounted on MirrIR slides.

6. Dewaxed FFPP sections: 10 μm waxed sections were immersed in a series of baths consisting of two baths of xylene (BDH, Dorset, UK) for five minutes and four minutes respectively, two baths of Ethanol Absolut (Merck, Dorset, UK) for three minutes and two minutes and a final bath of Industrial Methylated Spirits 95% (Lennox, Dublin, Ireland) for one minute.

Each category investigation was carried out on 3 parallel tissue pieces or sections within the category and a minimum of 10 spectra were recorded from each using both IR and Raman spectrometers. A representative selection is shown.

3.3 Results

Figure 3.1 compares the Raman spectra of fresh tissue and frozen tissue. Both tissues produce good spectra. There is an overall deterioration in the spectrum of the frozen tissue section compared to the fresh tissue. The most noticeable differences between the fresh and frozen tissue is the reduction in the intensity of the peaks at 1002 cm^{-1} (CC aromatic ring stretching), 1447 cm^{-1} (CH_2 bending mode of proteins and lipids) and 1637 cm^{-1} (Amide I band). The frozen tissue also displays an additional contribution at 1493 cm^{-1}.

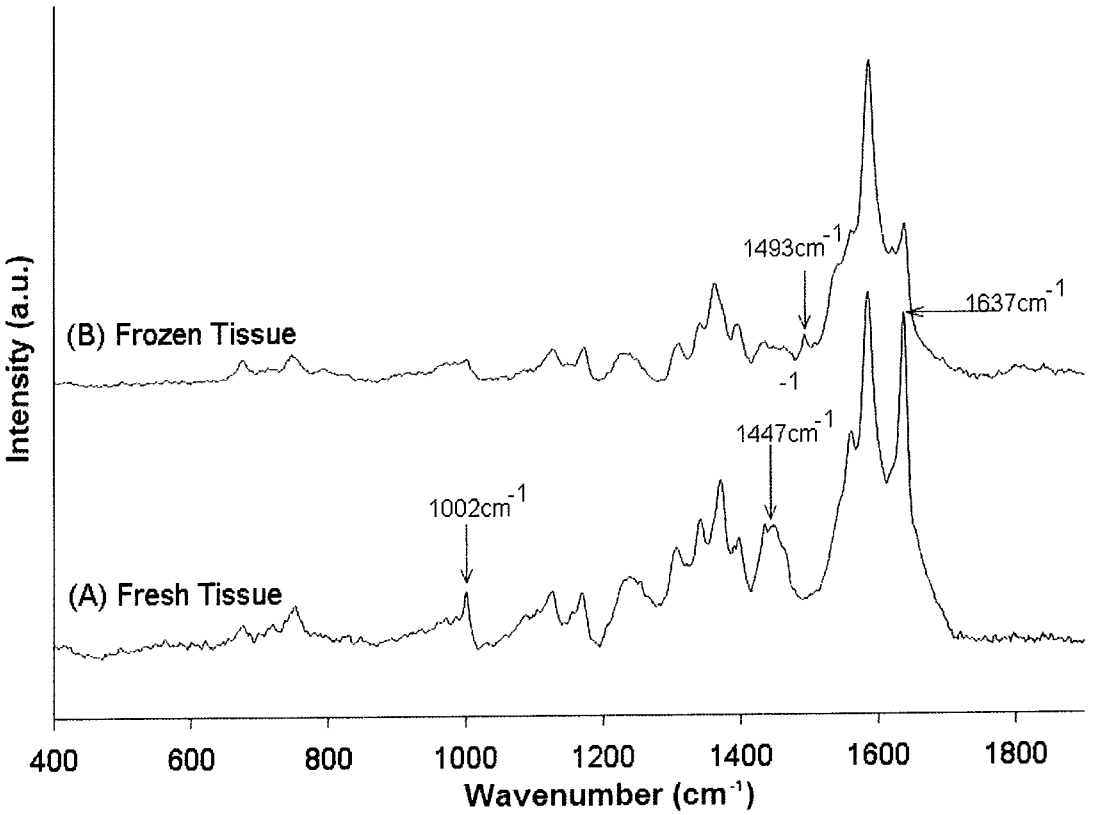


Figure 3.1 Raman spectra of (A) fresh tissue compared to (B) frozen tissue section

Interestingly the Raman spectra recorded from parenchyma tissue differ from those seen from a variety of different tissue types examined previously (Figure 2.16). Parenchymal tissue by it's vary nature is rich in blood and blood compounds. Presence of residual blood can be identified by the presence of heme groups (Puppels G.J. - personal communication) and these heme groups may be responsible for the unusual spectral properties of parenchymal tissue.

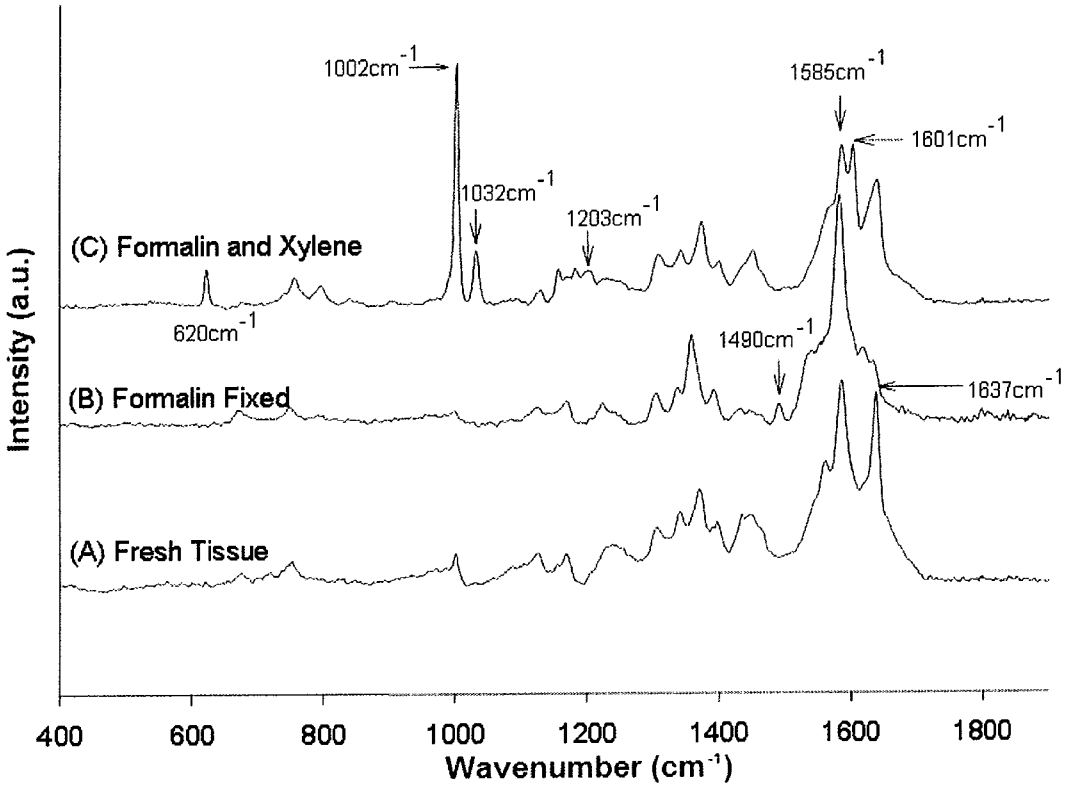


Figure 3.2 Raman spectra of (A) fresh tissue, (B) formalin fixed tissue and (C) tissue fixed and soaked in xylene

Figure 3.2 compares the Raman spectra of fresh tissue with formalin fixed tissue and tissue fixed and soaked in xylene. The most obvious change in the tissue with the fixation in formalin is the dramatic reduction in the intensity of the Amide I band (1637 cm^{-1}) and the addition of the peak at 1490 cm^{-1}. Soaking in xylene prior to wax embedding creates significant spectral differences in the tissue. Strong peaks appear at 620, 1002, 1032 and 1601 cm^{-1} (CC twist aromatic ring, CC stretching aromatic ring, CC skeletal stretch and CC in plane bending respectively). The addition of the small peak at 1203 cm^{-1} is attributed to the $C\text{-}C_6H_5$ stretch mode. There is also a reduction in the 1585 cm^{-1} band (C=C olefinic stretch) and the Amide I band has also reappeared at 1637 cm^{-1}.

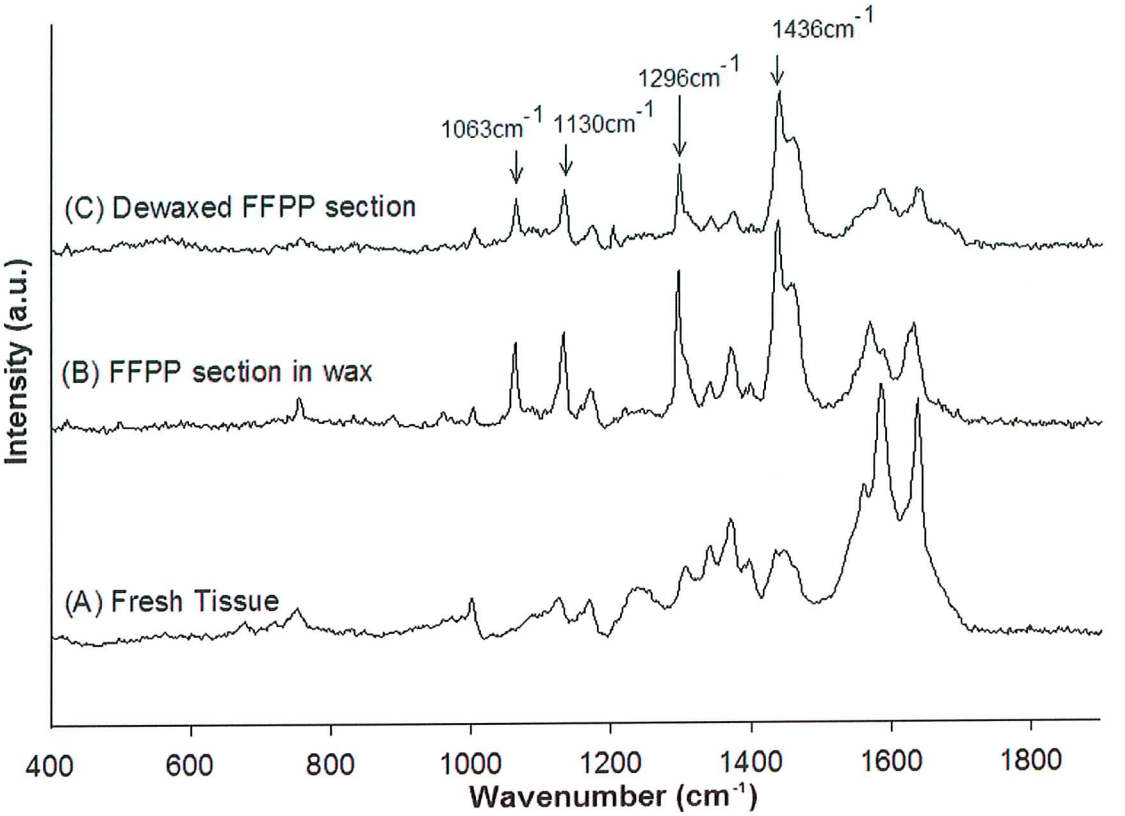


Figure 3.3 Raman spectra of (A) fresh tissue, (B) FFPP section in wax and (C) dewaxed FFPP section using xylene

Figure 3.3 compares the Raman spectra of fresh tissue with FFPP sections in wax and dewaxed FFPP sections using xylene as a dewaxing agent. In the FFPP section in wax contributions are clearly seen at 1063, 1130, 1296 and 1436 cm^{-1} (CC skeletal stretch random conformation, CC skeletal stretch *trans* conformation, CH_2 deformation and CH_2 scissoring respectively). A Raman spectrum of paraffin wax is shown in Chapter 4 (Figure 4.1) It is clear from the spectrum of the dewaxed FFPP section that all the wax is not being removed and residual wax peaks (at the same wavenumbers as outlined above) are present in the tissue after dewaxing. Also, as seen in the previous processing steps, there is

an overall reduction in signal from the biological material after freezing and formalin fixation.

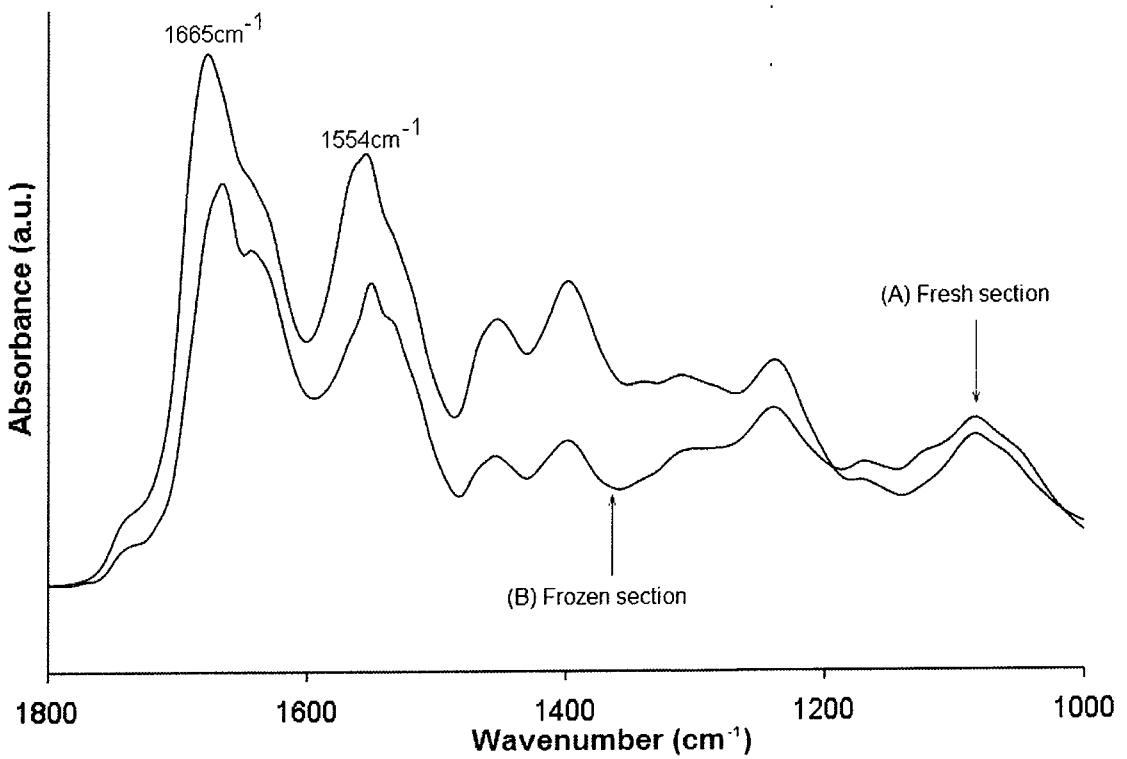


Figure 3.4 FTIR spectra of (A) fresh and (B) frozen tissue

Figure 3.4 examines the FTIR spectra of fresh and frozen tissue. The frozen tissue displays an overall reduction in intensity. Also, the Amide I and II bands (1673 and 1554 cm^{-1} respectively) appear shifted indicating possible changes in protein conformation in the frozen tissue section.

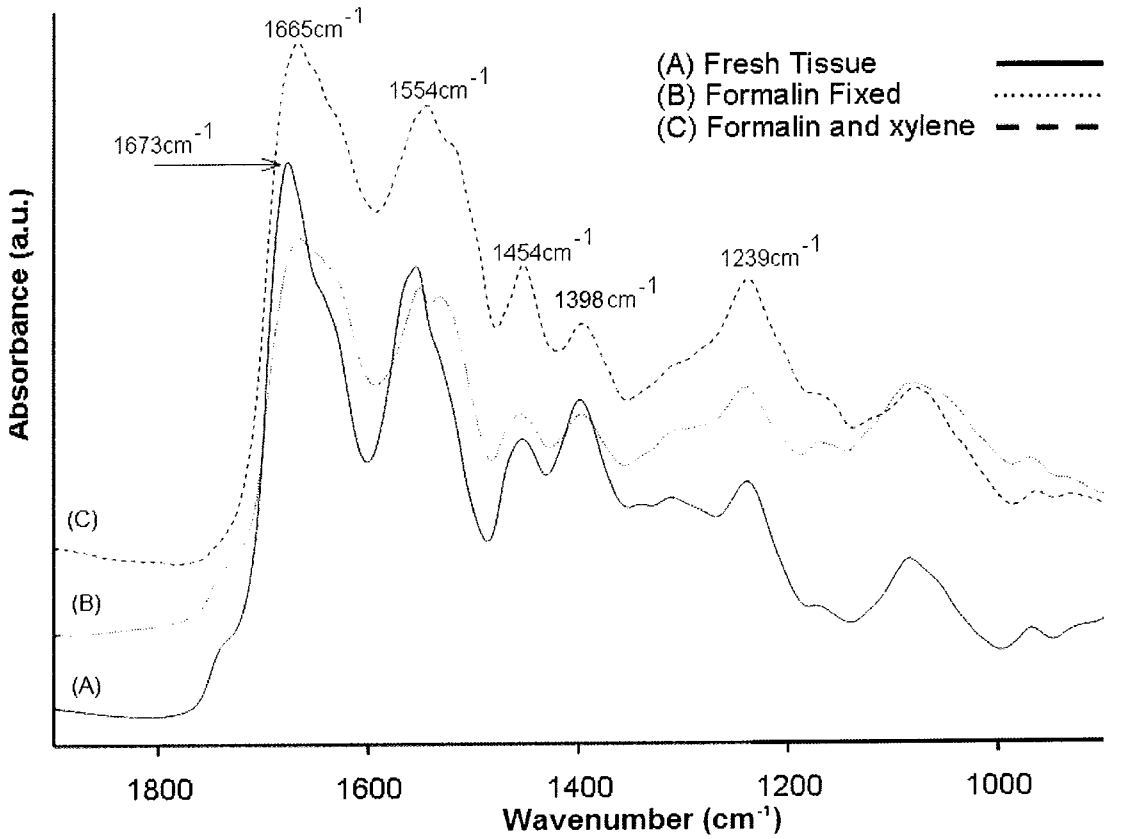


Figure 3.5 FTIR spectra of (A) fresh tissue, (B) formalin fixed tissue and (C) tissue fixed and soaked in xylene before wax embedding

Figure 3.5 examines the effect of formalin fixation and formalin and xylene processing before wax embedding using FTIR spectroscopy. There is an apparent 10 cm^{-1} shift in the Amide I and II bands after formalin fixation. The reduction in intensity of the Amide I band seen using Raman spectroscopy is mirrored in the FTIR spectra. Xylene results in an increase in intensity of the CH_2 scissoring (1454 cm^{-1}) and Amide III vibrations (1239 cm^{-1}). A reduction in the intensity of the 1398 cm^{-1} band (C=O symmetric stretch) is observed after fixation and xylene processing.

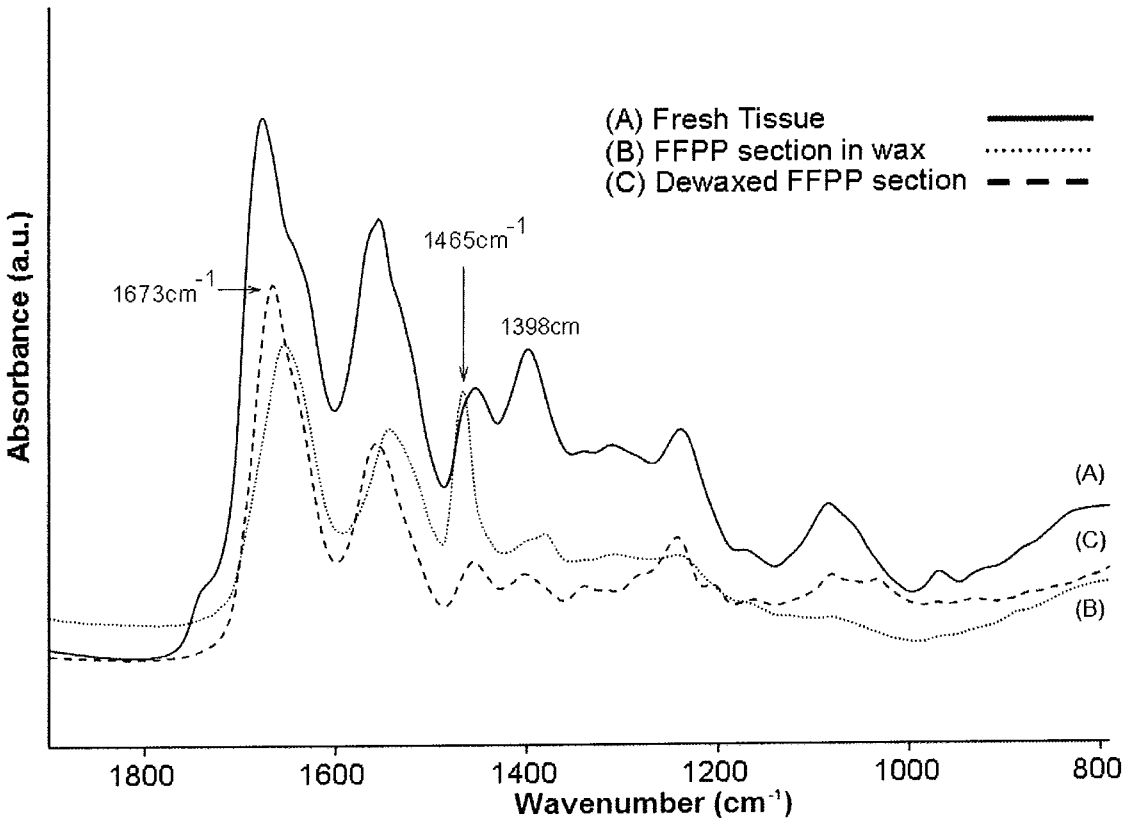


Figure 3.6 FTIR spectra of (A) fresh tissue, with (B) wax embedded tissue and (C) dewaxed FFPP section

Figure 3.6 compares the FTIR spectra of fresh tissue with wax embedded tissue and dewaxed FFPP sections. The wax peak is clearly visible in the waxed FFPP sections at 1465 cm^{-1} (CH_2 scissoring). The Amide I band has shifted back to its original position (1673 cm^{-1}) after dewaxing. Again there is a loss in signal intensity after dewaxing.

3.4 Conclusion

The Raman spectrum of frozen tissue compared to fresh tissue displayed a reduction in intensity of the 1002 cm^{-1} (CC aromatic ring stretching), 1447 cm^{-1} (CH_2 bending mode of proteins and lipids) and 1637 cm^{-1} (Amide I) bands. This overall reduction in Raman intensities was also observed by Huang *et al.* (2003). A new peak was also observed at 1493 cm^{-1}. In order to eliminate the mounting medium OCT as a possible contributing factor, Raman spectra were recorded from OCT. No Raman signal was observed at 1493 cm^{-1} and therefore OCT was not the cause of the additional peak. The effects of freezing on cells has been well documented in the cryogenics field and it is well known that a reduction in temperature can result in depolymerisation of the cellular cytoskeleton (Watson et al. 1987). The cellular cytoskeleton is composed of different types of protein fibres. Any depolymerisation of these proteins would result in unravelling of the secondary structure and hence an increase in the NH_3^+ and COO^- vibrations. The Raman frequencies for vibrations involving these groups occur between 1485-1550 cm^{-1} and 1560-1600 cm^{-1} respectively and both increase after the tissue had been frozen. Thus the appearance of the band at 1493 cm^{-1} can be attributed to an increased contribution from NH_3^+ deformation, as a result of protein structural changes on freezing tissues.

Fixation in formalin also produced a reduction in intensity of the Amide I band at 1637 cm^{-1}. Although all the effects of formalin fixation are not completely understood, the general principles are known. Aldehydes in formalin form cross-links between proteins creating a gel, thus retaining cellular constituents in their in-vivo relationship to each other.

Soluble proteins become fixed to structural proteins. The majority of cross-links are formed between the nitrogen atom of lysine and the nitrogen atom of a peptide linkage (Fig. 3.7).

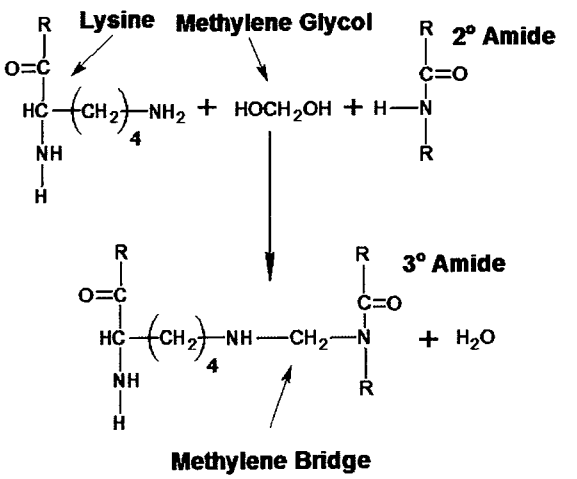


Figure 3.7 Scheme illustrating the formation of a methylene bridge during formalin fixation

This cross-link (methylene bridge) alters the amide from a secondary amide to a tertiary amide, which in turn alters the frequency of the C=O vibration, which could explain the loss of the band at 1637 cm^{-1}. The new peak at 1490 cm^{-1} may be due to protein unravelling, resulting in the increased activity of the NH_3^+ deformation similar to that seen after tissue freezing. Alternatively, a study by Wojciechowski *et al.* (2003) found a distinct Raman band at 1491 cm^{-1} which was assigned to the C-N stretching vibration coupled with the in-plane C-H bending mode in amine radical cations. Such coupled vibrations may be present in the methylene bridging system proposed to form after formalin fixation (Fig. 3.7). It can be assumed that there is no residual formalin due to the fact that other characteristic formalin peaks at 907 cm^{-1} and 1041 cm^{-1} (Huang *et al.* 2003) are not present and any excess formalin has been removed by rinsing in PBS.

Following xylene processing, xylene contributions were observed at 620, 1002, 1032, 1203 and 1601 cm^{-1}. These CC ring vibrations are to be expected given the aromatic structure of xylene. The absence of the 1490 cm^{-1} band associated with the presence methylene bridges suggested above and the re-appearance of the Amide I band (1637 cm^{-1}) suggests that xylene has reversed some of the cross-links created with formalin fixation and the tertiary amide band has reverted to a secondary amide as seen in fresh tissue. The xylene signature peaks were not present in the dewaxed tissue so it may be assumed that all xylene was being removed after dewaxing. However, it is clear from the Raman spectrum of dewaxed FFPP sections that all the wax was not removed. Strong wax contributions were seen at 1063, 1130, 1296 and 1436 cm^{-1}, so residual wax is an issue when dealing with FFPP sections.

Using FTIR spectroscopy, fresh tissue provided the most detailed spectra, whereas freezing resulted in loss of intensity and a shift in the Amide I and II bands. This shifting may result from the depolymerisation outlined previously. Formalin fixation produced a shift of 10 cm^{-1} in the amide I and II bands, which corroborates the findings of Pleshko *et al.* (1992) who found the amide bands shifting when fixed in ethanol. Freezing and formalin fixation resulted in a large reduction in the intensity of the 1398 cm^{-1} band (C=O symmetric stretch), which corroborates the Raman findings. Xylene resulted in an increase in the intensity of the 1239 and 1454 cm^{-1} bands, but did not display the dramatic spectral differences seen using Raman spectroscopy. The FTIR signature contribution of wax at 1465 cm^{-1} was also visible. After dewaxing there was a reduction in intensity of the bands. The Amide I band shifted back to its original position of 1673cm^{-1} as seen in fresh tissue

and the wax signature at 1465 cm^{-1} has been reduced. However it was very difficult to say whether or not the wax had been removed using FTIR spectroscopy. Again, in this instance Raman spectroscopy provides more detailed biochemical information. Therefore the 1454 cm^{-1} band (CH_3 asymmetric deformation) seen in fresh tissue and dewaxed FFPP sections, could easily be contaminated with contributions from wax occurring at 1465 cm^{-1}. For this reason it is not advisable to use this peak in a diagnostic capacity.

The reduction in overall signal intensity after processing seen in both Raman and FTIR remains unexplained. The spectra of fresh, formalin fixed tissue were recorded from tissue pieces mounted on MirrIR slides and therefore would inherently have variations in sample thickness. However any variation in sample thickness would not have an effect on the intensity of the Raman spectra. This is due to the fact that Raman spectroscopy is a scattering effect, with a penetration depth of $\approx$2 μm, far below the minimum sample thickness of $\approx$10 μm. Sample thickness would affect the overall FTIR signal intensity, due to the double transmission nature of the measurement. But the fact that this signal reduction was also seen in the Raman spectra confirms that the effect results from more than sample thickness variation. It is possible, however, that differences in sample density between tissue sections and excised tissue pieces may account for the differences in overall signal intensity observed between samples. Although dewaxed FFPP sections display a reduction in intensity, all of the bands in fresh tissue are present and many are diagnostically useful.

Generally speaking it was found that Raman spectroscopy produced more detailed spectra and hence appears more sensitive to changes in tissue composition than FTIR spectroscopy.

Naturally, Raman spectroscopy is intrinsically more sensitive to some bands than FTIR and *vice-versa*. However it is the higher spectral and spatial resolution that results in the more detailed spectra obtained using Raman spectroscopy. The spectral resolution of the Raman spectrometer (2 cm^{-1}) is superior to the 8 cm^{-1} resolution used with the FTIR spectrometer. Also, the much larger sampling area of the FTIR spectrometer (although beneficial for sampling large areas) will also have the effect of averaging over many cells, broadening bands with the loss of spectral information.

In summary, although bearing a close resemblance to fresh tissue, processed tissue has undergone many biochemical changes as shown. This systematic study has provided spectra at each of the processing stages involved in dewaxed FFPP sections. This gives an insight into the biochemical changes resulting from each process. It has highlighted changes the tissue has undergone that would otherwise be unseen when comparing fresh with dewaxed FFPP sections. This study has confirmed that dewaxed FFPP sections can be diagnostically useful as they retain sufficient biochemical similarity to fresh tissue. However, researchers in this field should be aware of the spectral contributions affected by processing, particularly when utilising automated analysis techniques. Residual wax contributions were found to be the most problematic issue when using dewaxed FFPP sections. Further investigations to improve the level of wax removal are the subject of Chapter 4.

Chapter 4

Optimisation of Dewaxing Protocol

4.1 Introduction

As seen in the previous chapter, residual wax is an issue when examining dewaxed FFPP tissue sections using vibrational spectroscopy. The wax signature peaks are most apparent using Raman spectroscopy as opposed to FTIR and therefore Raman spectroscopy was used to evaluate the protocol for wax removal in tissue sections. The work presented in this chapter has been published in the Journal of Histochemistry and Cytochemistry (Ó Faoláin *et al.* 2005b)

Microscopical examination of tissue requires a slice of the tissue thin enough to transmit light (frequently 3-20μm). In the preparation of such thin slices (section cutting or microtomy) the tissue must undergo preparatory treatment before being sectioned, resulting in impregnation of the specimen with an embedding medium which provides support and a suitable consistency for microtomy (Bancroft *et al.* 2002). This preparatory treatment is known as tissue processing. The most commonly used embedding medium is paraffin wax (Drury *et al.* 1980). The paraffin wax method involves the dehydration of tissues, clearing in organic solvent and their impregnation with molten paraffin wax, which is subsequently allowed to solidify by cooling. Cutting qualities are good, the blocks are durable and can be stored long term without any deterioration. There is no doubt that if tissue is intended for microscopical examination at some future date it keeps better in the form of a paraffin block than in any fixative or preservative (Drury *et al.* 1980). Also evaluation of antigens and nucleic acids is possible in tissues stored in this medium for long periods (Bonin *et al.* 2003, Schurter *et al.* 2002, Chan *et al.* 2001). For all of the reasons above, it is the most

commonly used embedding medium in both normal and pathological histology, allowing both diagnostic and retrospective studies.

FFPP tissue sections are used extensively for immunohistochemical detection of normal and tumour cell markers. A major step forward was made in the 1990's with the discovery that some antigens previously unreactive in FFPP tissue, even after protease treatment, could be "retrieved" by heating sections in a microwave oven. Originally this was carried out in a solution of rather toxic heavy metal salts (Shi *et al.* 1991). Subsequently these salts were replaced with buffers such as citrate buffer at pH 6.0 (Cattoretti *et al.* 1993). It was shown that heat rather than microwaves *per se,* is important in the retrieval process. Boiling the section in a pressure cooker (Norton *et al.* 1994) or autoclaving them (Bankfalvi *et al.* 1994) in the buffer solution achieved the same effect (Polak *et al.* 1997). A more recent dewaxing protocol involves the application of a reagent for simultaneous dewaxing and antigen unmasking, namely Declere® and Trilogy™ (www.cellmarque.com). These protocols are used extensively for immunohistochemistry and *in situ* hybridisation techniques.

To examine the biochemical structure of the FFPP tissue using Raman spectroscopy, it is desirable that the sample be as close to its original/*in vivo* state as possible. This requires the removal of the paraffin wax and rehydration to the aqueous phase. In this present study multiple methods of dewaxing were examined and the efficacy of each assessed using Raman spectroscopy.

This study compared the effectiveness of the most commonly used dewaxing agents, namely, xylene and Histoclear™, as well as heat mediated antigen retrieval (HMAR) using

xylene followed by a citrate buffer and HMAR using Trilogy™ alone. The effectiveness of hexane (known for its uses as an industrial cleaning and degreasing solvent) as a possible dewaxing agent was also examined (personal communication S.M. Hewitt). The impact of wax removal on immunohistochemical staining was also addressed. Therefore samples were stained immunohistochemically after dewaxing with xylene and hexane and compared.

4.2 Materials and Methods

4.2.1 Instrumentation

The Instruments S.A. Labram Raman spectroscopic confocal microscope was used, with an Argon Ion laser operating at a wavelength of 514.5nm and a x50 objective lens. The laser power at the sample was measured and found to be ≈6.5±0.05mW, focused to a spot size of 2μm at the tissue surface. The scattered Raman signal was integrated for 60 to 150 seconds and measured over a spectral range of 400 to 1900 cm^{-1} with respect to the excitation frequency. Once acquired each spectrum was baseline corrected, dark current subtracted and lightly filtered using a third order linear model to improve clarity.

4.2.2 Sample Preparation for Raman Study

All cervical FFPP sections were obtained from the National Maternity Hospital, Holles St., Dublin. All specimens were automatically processed to wax as follows:

(a) vacuum fixed in 10% buffered formal saline histo-grade pH 6.8-7.2 (J.T. Baker, Deventer, The Netherlands) heated to 35°C (b) vacuum dehydration in industrial methylated spirit IMS T100 (Lennox, Dublin, Ireland) heated to 35°C (c) vacuum clearing in xylene (Serosep, Limerick, Ireland) heated to 35°C (d) vacuum impregnation with Tissue Tek III Embedding Wax with polymer added (Sakura, Zoeterwoude, The Netherlands) and heated to 59°C.

Three parallel FFPP cervical sections were dewaxed using each of the five protocols outlined below.

A. Xylene:

After the wax impregnation, tissue was embedded and sliced into 10μm sections using a microtome, mounted on glass slides and dried. The unstained samples were immersed in a series of baths consisting of two baths of xylene (BDH, Dorset, UK) for five minutes and four minutes respectively, two baths of Ethanol Absolut (Merck, Dorset, UK) for three minutes and two minutes and a final bath of Industrial Methylated Spirits 95% (Lennox Dublin, Ireland) for one minute.

B. Histoclear™:

The same procedure was used for dewaxing using General Purpose Grade Histoclear™ (Fisher Scientific, Loughborough, UK), where xylene was substituted with Histoclear™ in protocol A.

C. Hexane:

Again, the same procedure was used for dewaxing using Hexane (BDH, Dorset, UK), where xylene was substituted with Hexane in protocol A.

Each of the sections examined was put through the dewaxing procedure outlined four successive times. Finally each of the specimens was left sitting in a bath of xylene, Histoclear^{TM}, or hexane for 18 hrs.

Spectra were recorded between each of the successive cycles and after immersion in reagents for 18 hours. Raman spectra were taken from 10 random points from each of the 3 sections, for each protocol. All spectra were recorded from normal ectocervical squamous epithelial cells.

D. Xylene dewaxing and HMAR in citrate buffer:

To examine the effect of the heat mediated antigen retrieval technique (HMAR) using xylene and a citrate buffer on the wax content, sections were processed according to the following HMAR protocol. Sections (mounted on positively charged slides) were passed through two changes of xylene for 6 minutes each and three changes of spirit for 3 minutes each. Slides were then submerged in 600ml of 0.1M citrate buffer (pH 6) and placed in a pressure cooker for 20 mins.

E. Simultaneous dewaxing and HMAR in Trilogy^{TM}:

Two staining dishes were filled with approximately 200ml of Trilogy^{TM} (Cell Marque Corporation, Arkansas, USA) and the sections (mounted on positively charged slides) were

submerged in the first staining dish. Both dishes were placed in a pressure cooker for 8 mins on the high-pressure setting. After 8 mins the pressure was released and the slides were transferred to the second dish (hot rinse) to soak in Trilogy™ for a further 10 mins. The slides were agitated and rinsed in de-ionised water.

Raman spectra were recorded from 10 random points of normal ectocervical squamous epithelial cells from each of the 3 HMAR treated sections.

Raman spectra were also recorded from paraffin wax sections, tissue sections prior to deparaffinisation and frozen sections not embedded in paraffin wax. All sections were air dried and examined spectroscopically.

4.2.3 Sample Preparation for Immunohistochemical Study

Sections were dewaxed for 18 hours at room temperature in hexane or xylene respectively. Sections were rehydrated to water through Ethanol Absolut (Merck, Dorset, UK) and Industrial Methylated Spirits 95% (Lennox Dublin, Ireland). The sections were placed in 0.1mM citrate buffer, pH 6 and microwaved for 12 minutes at 750W. They were kept in the hot solution for a further 20 minutes, washed in water, treated with 0.3% hydrogen peroxide in methanol for 10 min, then transferred to phosphate buffered saline (PBS), pH 7.2. The Vector Elite Avidin-Biotin immunoperoxidase kit (Vector Laboratories, Peterborough, U.K.) was employed for the immunoperoxidase method. Slides were treated with normal horse serum 1:50 for 10 min and then the monoclonal antibody to cytokeratin MNF 116 (Dakocytomation, Glostrup, Denmark) was applied at a dilution of 1:100 in PBS

for 60 min at room temperature. Sections were washed in PBS for 10 min and biotinylated secondary antibody (1:200) was applied for 15min. Following a wash in PBS for 10min the avidin-biotin complex solution (1:50) was applied for 15min, sections were washed in PBS for a final 10min and incubated in the chromogenic DAB substrate for 5 min (0.003% hydrogen peroxide and 0.06g/ml diaminobenzidene in PBS). A nuclear counterstain with haematoxylin was applied, the slides were dehydrated through alcohols, placed in xylene and coverslipped with DPX resin medium. After the immunohistochemical protocol, images of samples were recorded using a Nikon Eclipse E600 optical microscope equipped with RT Spot digital camera (Digital Instruments), with a x40 objective.

4.3 Results

4.3.1 Results of Raman Study

The spectrum of paraffin wax clearly shows distinctive Raman contributions at $1063cm^{-1}$, $1133cm^{-1}$, $1296cm^{-1}$ and $1441cm^{-1}$ (Fig. 4.1). These are attributed to C-C stretching and CH_2 and CH_3 deformation modes (Barry *et al.* 1992, Lakshmi *et al.* 2002), as would be expected given the straight chain hydrocarbon structure of wax (Fig 4.2). A small contribution at $1004cm^{-1}$ (C-C aromatic ring) is attributed to the small percentage of cycloparaffins contained within the wax.

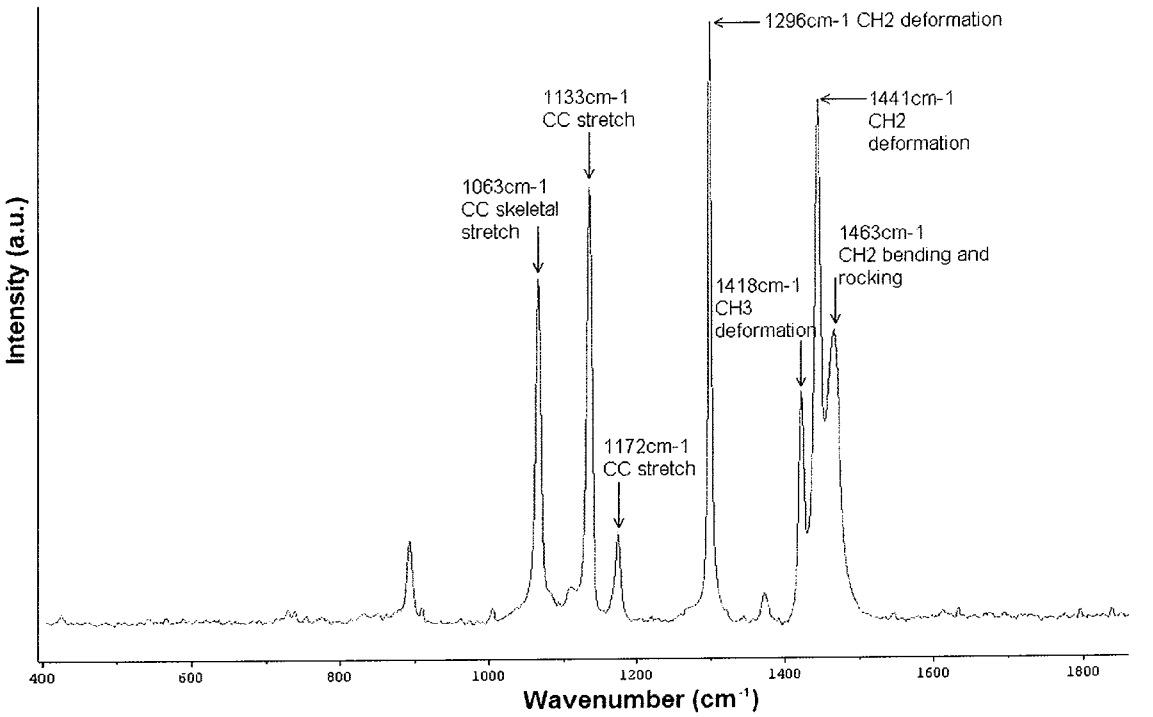


Figure 4.1 Raman spectrum of paraffin wax

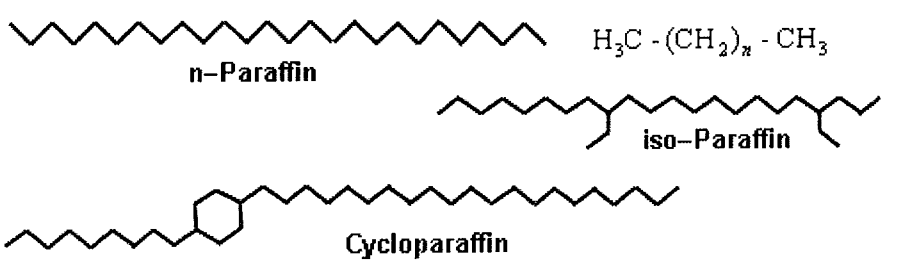


Figure 4.2 Components of paraffin wax

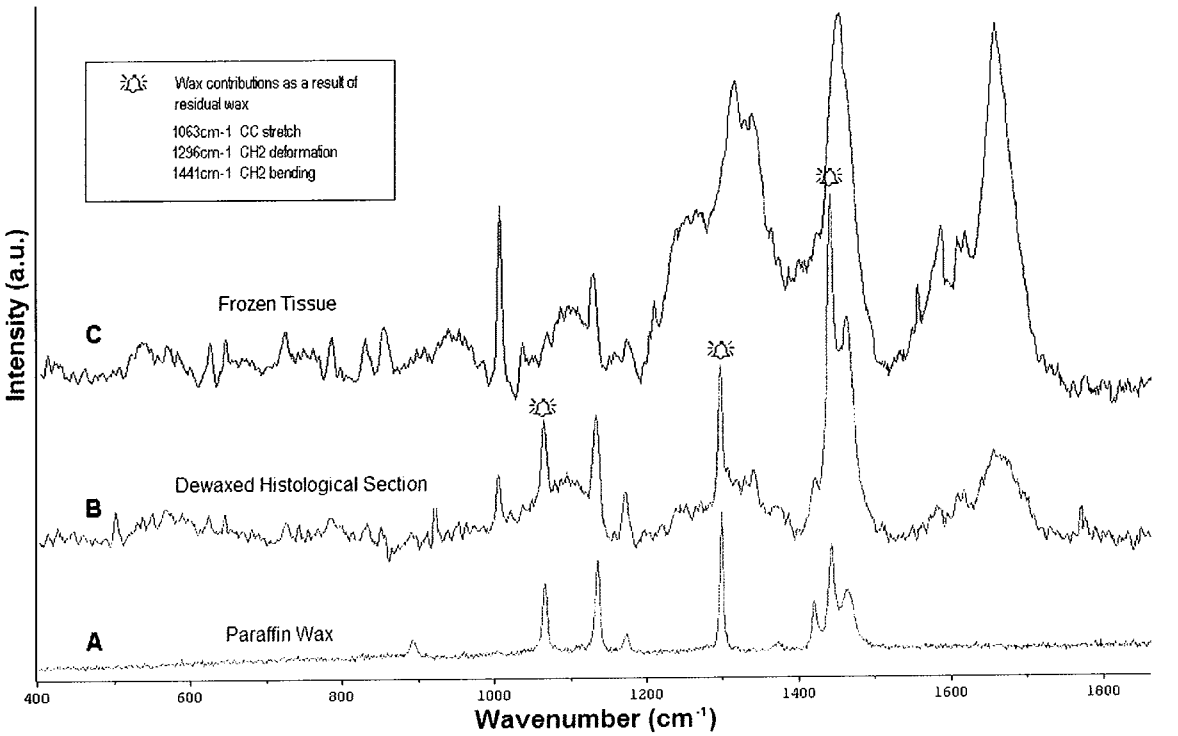


Figure 4.3 Raman spectra comparing paraffin wax, dewaxed FFPP section and frozen tissue

Figure 4.3 compares the Raman spectra of A) paraffin wax, B) an ineffectively dewaxed histological section (using xylene) and C) a frozen tissue section that has not been embedded in wax. It demonstrates that the contributions from the paraffin wax seen at $1062cm^{-1}$, $1296cm^{-1}$ and $1441cm^{-1}$ are visible in the dewaxed FFPP section (Fig. 4.3B) and that they are not present in unprocessed tissue (Fig. 4.3C). The C-C stretch band at $1133cm^{-1}$ appears in both paraffin wax and biological samples and for this reason has not been highlighted as an exclusive paraffin contribution. Many additional peaks are present in both the spectra of the dewaxed FFPP section and the frozen section. These peaks correspond to vibrations of bonds emanating from the biological material itself. The spectra are dominated by peaks in the regions of 1330 cm^{-1}, 1450 cm^{-1} and 1659 cm^{-1}, corresponding to

CH_2 and CH_3 twisting and wagging, CH_2 bending mode of proteins and the amide I band (C=O stretching) of proteins respectively. Other peaks include 622 cm^{-1} (CS stretch), 644 cm^{-1} (CS stretch; Amide IV), 725 cm^{-1} (=CH in plane bending), 829 cm^{-1} (CCH aliphatic deformation), 853 cm^{-1} (CCH aromatic deformation), 1172 cm^{-1} (CC stretch), 1583 cm^{-1} (C=C bending), 1601 cm^{-1} (C=C in plane bending) and 1620 cm^{-1} (C=C stretching) (Stone *et. al* 2000, Barry *et. al* 1992).

As the units associated with spectroscopic intensity are arbitrary units (a.u.), it is only the relative intensity to the baseline of each spectrum that is relevant. Hence, the spectra have been offset in order to facilitate clarity. This method of display has been used throughout.

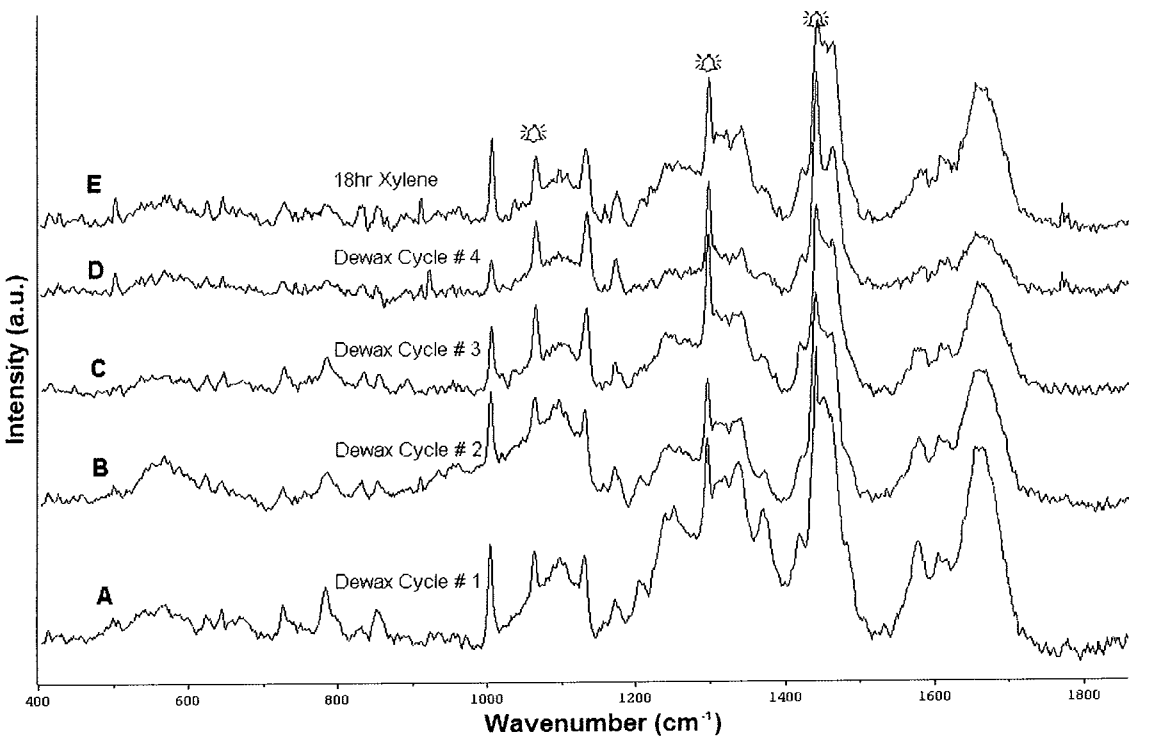


Figure 4.4 Raman spectra after each subsequent dewaxing cycle using xylene

The effectiveness of xylene (Fig. 4.4) and Histoclear™ (Fig. 4.5) on the wax content were examined and found to be ineffective at complete removal of wax. The Raman spectra after a single dewaxing cycle (Fig. 4.4 A) and subsequent dewaxing cycles (Fig. 4.4 B, C, D) showed the signature wax contributions at 1062 cm^{-1}, 1296 cm^{-1} and 1441 cm^{-1}. These contributions were not eliminated even after immersing slides for 18 hours in xylene (Fig. 4.4 E).

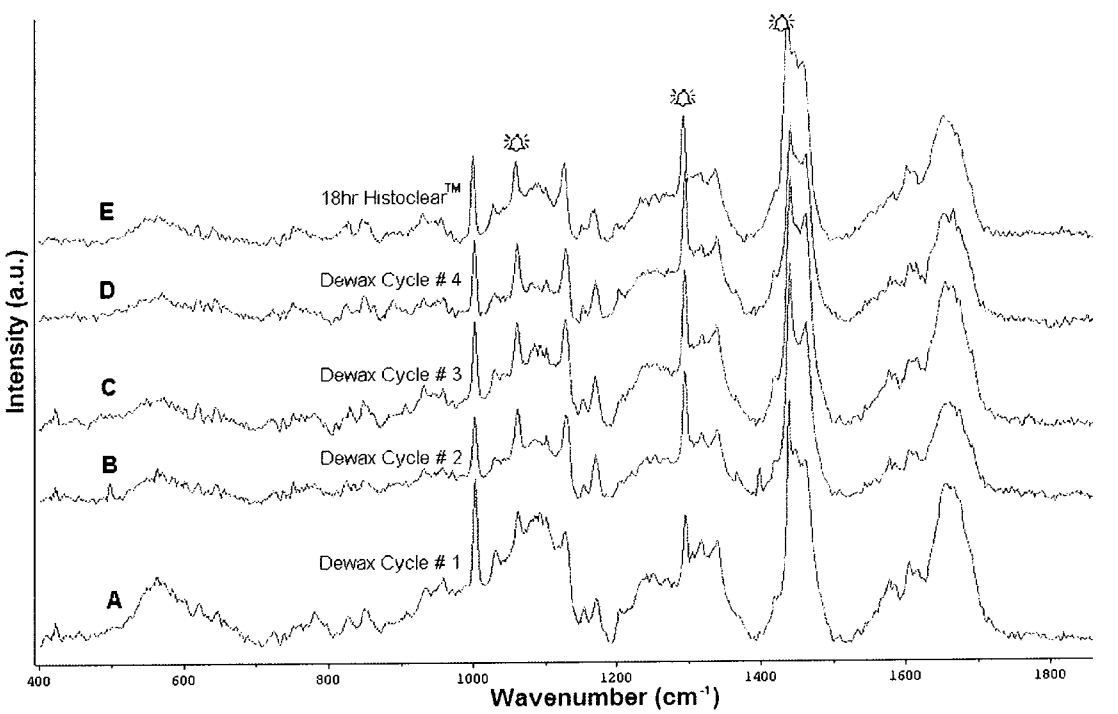


Figure 4.5 Raman spectra after each subsequent dewaxing cycle using Histoclear™

The same residual wax contributions were present in spectra after a single dewaxing cycle (Fig. 4.5 A), subsequent dewaxing cycles (Fig. 4.5 B, C, D) and immersion in Histoclear™ for 18 hours (Fig. 4.5 E).

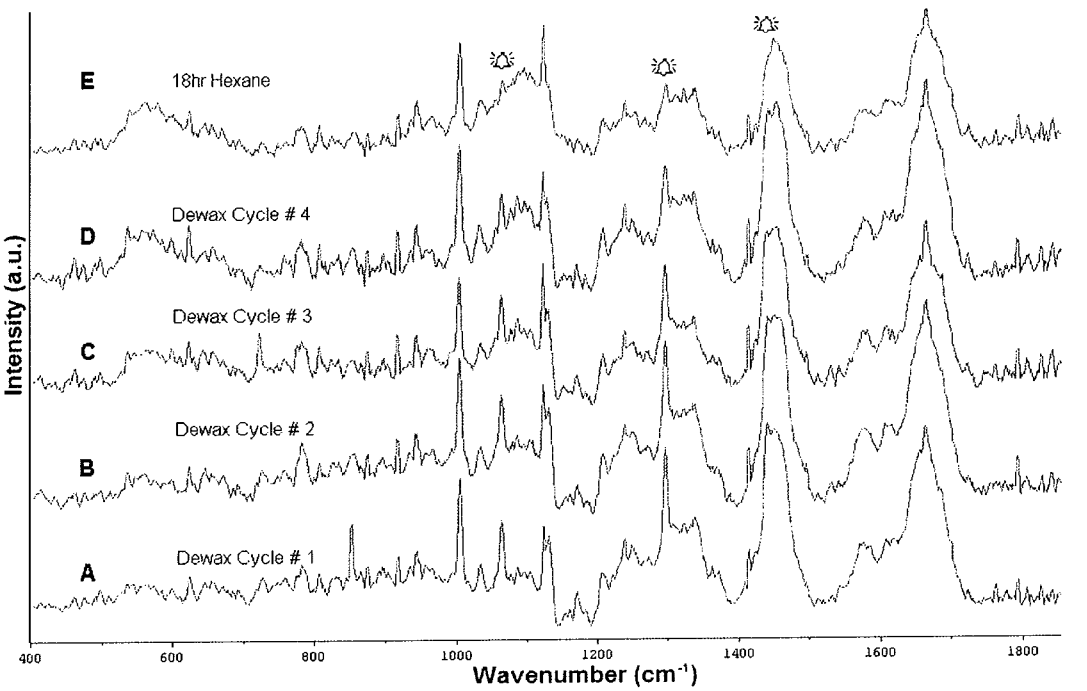


Figure 4.6 Raman spectra after each subsequent dewaxing cycle using hexane

Hexane, however (Fig. 4.6), proved to be much more effective at wax removal than both xylene and Histoclear™. The wax content reduced after each subsequent cycle (Fig. 4.6 A, B, C, D) and had almost been completely removed after immersion in hexane for 18 hours (Fig. 4.6 E).

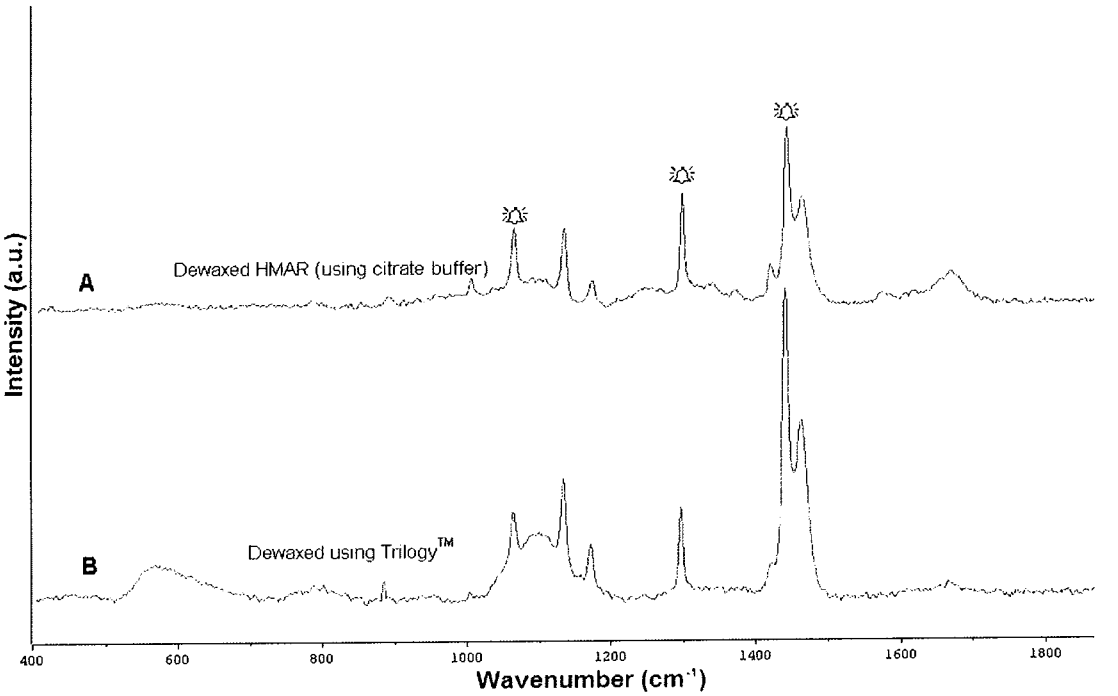


Figure 4.7 Raman spectra comparing dewaxing HMAR (using xylene and citrate buffer) and Trilogy™

Following dewaxing with xylene and HMAR using citrate buffer, residual wax contributions were observed at 1062cm^{-1}, 1296cm^{-1} and 1441cm^{-1} (Fig. 4.7 A). Following simultaneous dewaxing and HMAR with Trilogy™ the same residual wax contributions were also observed (Fig. 4.7 B). Due to the high pressure and temperature involved in the pressure cooker technique, the tissue begins to degrade after multiple cycles. This tissue degradation is evident in the Raman spectra in Figure 4.7, where an overall deterioration in signal intensity compared to the signals measured in Figures 4.3, 4.4, 4.5 and 4.6 can be observed. However multiple cycle investigations were carried out and residual wax remained (results not shown).

4.3.2 Results of Immunohistochemical Study

The impact of wax removal on immunostaining was evaluated by directly comparing parallel sections dewaxed using xylene and hexane after immersion for 18 hours. The immunostaining protocol was applied to both slides, where brown staining indicated immunopositivity (antigen/antibody binding). Subsequently images were recorded from the same area of both slides and were compared. Fig. 4.8 shows images taken from both slides after dewaxing with xylene and hexane (Fig. 4.8 A, B respectively).

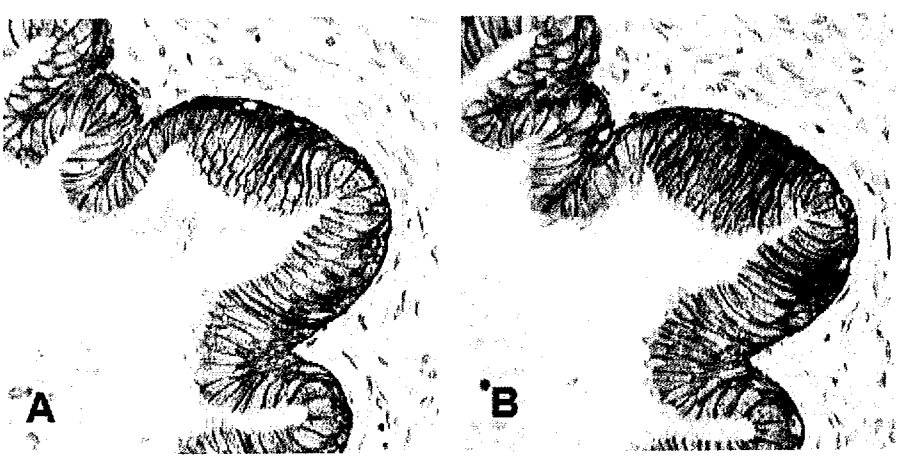


Figure 4.8 Images of samples after immunostaining

Qualitatively, it is clear from Figure 4.8 that there is a stronger positivity in Figure 4.8 B (dewaxed using hexane), than 4.8A (dewaxed using xylene). To quantify this increase in positivity, both images were converted to greyscale and the overall intensity of the stained region was determined using the ImageJ analysis programme (National Institute of Health USA). The intensity quantifies the amount of immunostain uptake and hence can be used to quantify the overall degree of positivity. The overall intensity after dewaxing using xylene

was 124.99±5.03 a.u., whereas the value measured from the slide dewaxed using hexane was 97.46±1.52 a.u. This reduction in intensity quantifies the increase in positivity, which has been improved by 28% when dewaxed using hexane as opposed to xylene.

4.4 Conclusion

The present study investigated five clearing agents / processes (xylene, Histoclear™, hexane, HMAR (using xylene and citrate buffer) and Trilogy™). Raman spectroscopy clearly demonstrated that peaks corresponding to the presence of wax were present in all samples examined and that wax is not currently being removed with existing protocols. Hexane was shown to be more effective at wax removal than all other agents examined. This increased level of wax removal resulted in an approximate 28% improvement in immunohistochemical staining compared to a parallel section dewaxed using xylene.

Xylene is the most commonly used solvent, due to its rapid and supposedly efficient dewaxing for common histological, histochemical and immunohistochemical procedures. It is a highly toxic, flammable substance and must be handled with care. Histoclear™ on the other hand is non-toxic, non-flammable and biodegradable, but is a less efficient clearing agent and potentially causes the haematoxylin stain to fade (Culling *et al.* 1985). The HMAR technique uses xylene as a clearing agent and reverses formaldehyde fixation effects using citrate buffer in a pressure cooker (MacIntyre 2001). Trilogy™, on the other hand, is a novel product that combines deparaffinisation, rehydration and unmasking of antigens during pressure cooking. However Trilogy™ is under evaluation for application in

diagnostic immunohistochemistry and this is the first report comparing the dewaxing efficacy of Trilogy™ to the more standard reagents using Raman spectroscopy. Hexane is a colourless liquid with a slight odour and like many solvents is highly flammable. It should be handled with care and used only in a fume hood. The US Department of Health and Human Services, International Agency for Research on Cancer and the US Environmental Protection Agency have not classified hexane for carcinogenicity. This study is also the first to examine the dewaxing properties of hexane and directly compare it to all commonly used dewaxing agents.

In Chapter 3, it was noticed that the standard dewaxing procedure used on FFPP tissue did not completely remove the wax as previously supposed. A quantity of residual wax remained. From a spectroscopic perspective, with a view to detection of biochemical changes in the composition of FFPP sections, residual wax is problematic. Its spectral contributions add a substantial degree of variation to the overall spectrum. Contributions at $1062cm^{-1}$, $1296cm^{-1}$ and $1441cm^{-1}$, corresponding to CC skeletal stretch, CH_2 deformation and CH_2 bending respectively, mask contributions from these modes emanating from the tissue. Thus, potential biochemical changes in these regions may be rendered useless. Also, caution must be exercised when using automated analysis techniques while there are random variables in the datasets.

Although it is possible to use automated data extraction techniques to remove the effects of wax on the tissue spectrum (Tfyali *et al.* 2005), it is preferable to try to ensure that wax is being effectively removed. Ineffective removal of wax from sections can cause birefringence of cell nuclei (Nedzel *et al.* 1951). In addition, failure to completely remove

the wax from the sections will result in impairment of staining. This can result in the "Pink Disease" artifact (Vlachos *et al.* 1968). The artifact, which is most noticeable in lymphoid and epithelial tissue, is an extremely patchy distribution of stains and results in the loss of distinction of nuclear margins (Drury *et al.* 1980). There are also implications for antigen demonstration by immunohistochemistry. It is only in recent times that immunohistochemistry has been carried out on paraffin fixed sections. It is clear from the results of both HMAR and Trilogy™ investigations that all the wax is not being removed.

This research is the first study to report that a low level of residual wax is present in solvent-treated paraffin wax sections. The majority of histochemical and immunohistochemical methods employed in histopathology laboratories utilise these solvents as their dewaxing agents. Many dyes stain more intensely on cryostat sections and some do not work well on paraffin sections, particularly metachromatic stains for carbohydrates (H. Lambkin, personal communication), thus the residual wax may be a factor contributing to these differences. Immunostaining has been optimised for paraffin sections with hundreds of antigens now detectable in these preparations. Heat-based antigen unmasking has been introduced since the 1990s and has contributed to a lowering of the antigen detection threshold in paraffin sections (Shi *et al.* 1991). However, the heat effects are considered to be related to reversing of the effects of formaldehyde cross-linking of proteins, rather than removal of wax.

This study has demonstrated that current dewaxing procedures are not completely effective. Hexane has been identified as a superior dewaxing solvent to xylene as well as Histoclear™ and Trilogy™. It has also been demonstrated that increased wax removal

using hexane results in better antigen/antibody binding and hence a stronger positivity. In addition, care should be taken when using FFPP sections for spectroscopic investigation of diseased tissue, particularly choosing peaks for automated analysis. This study recommends dewaxing using hexane to minimise wax contributions.

Chapter 5

FTIR spectroscopy of cervical tissue

5.1 Introduction

Previous chapters examined the components of human tissue, the spectral properties of tissue at each different stage of tissue processing and great improvements were made in the removal of residual wax from dewaxed FFPP tissue sections. This chapter aims to identify spectral markers of cancer as a much better understanding of the various processes the tissue has undergone has been gained. This should ensure that only valid spectral markers of cancer are identified and not spectral changes associated with processing.

The first papers on FTIR studies of exfoliated cervical cells and cervical tissues appeared in the early 1990's (Wong *et. al* 1991, Wong *et. al* 1993). Wong *et al.* found several spectral differences in malignant as compared to normal cells and tissues. The spectroscopic differences included changes in intensity of bands at 1303, 1244, 1155, 1082, 1047 and 1025 cm^{-1}; shifting of peaks at 1244, 1155 and 1082 cm^{-1}; and the appearance of a band at 970 cm^{-1}. The ratio of intensities of the band due to glycogen at 1025 cm^{-1} to the band due to nucleic acids at 1082 cm^{-1} was found to differ between malignant and normal cells. They also observed that spectra of dysplastic cells were intermediate in magnitude between the spectra of normal and malignant cells and that the frequency of the 1082 cm^{-1} band was not shifted to higher frequencies, as for malignant samples. It was concluded that in malignant cells and tissues, the amount of glycogen is dramatically decreased, significant increases occur due to the hydrogen bonding of phosphodiester groups of nucleic acids, a loss of C–OH groups of proteins takes place and a greater disorder occurs in the conformation and orientation of the methylene chains of membrane lipids. The glycogen level decreases

during carcinogenesis, so decreases in intensity of the bands at 1155 and 1023–1025 cm^{-1}, which are due to absorption of glycogen, are understandable.

Wong *et al.* (1995) examined the spectral properties of exfoliated cervical cells and tissues from the endocervix and ectocervix by FTIR and attenuated total reflectance (ATR) FTIR spectroscopy. The study examined the FTIR spectra of cervical columnar cells, the underlying connective tissue and squamous ectocervical cells. Differences in many spectral features between normal ectocervical squamous cells and normal endocervical mucin–filled columnar cells were evident. Ectocervical cells displayed strong glycogen bands at 1023 and 1156 cm^{-1}, while columnar cells showed strong nucleic acid bands at 1082cm^{-1} and 1238cm^{-1} assigned to the symmetric and asymmetric stretch modes of the phosphodiester groups. The spectrum of connective tissue showed an absence of glycogen. However, sharp bands were seen at 1031, 1205, 1283, 1318 and 1139cm^{-1}, bands not seen in the spectrum of exfoliated endocervical cells. The study recommends the use of the ATR/FTIR method, rather than transmission FTIR when recording from a thin-layer of cells such as the single-cell layer of the epithelial tissue in the endocervix. This avoids removing a single layer of cells from underlying tissue, thus eliminating the potential of contamination from several layers of tissue.

Yazdi *et al.* (1996) studied cervical samples from 133 women and concluded that changes in the IR spectra are also clearly observed in cervical cells from pre-invasive lesions and other conditions such as benign cellular changes. Although confirming that FTIR spectra of abnormal cervical samples are different from the spectra of normal samples, it was shown that it is the high-grade squamous intraepithelial lesion (HSIL) specimens (95% of samples)

that exhibited the most dramatic differences in the IR spectra and that only 54% of low-grade squamous intraepithelial lesion (LSIL) cases and 33% of atypical squamous cells of undetermined significance (ASCUS) cases exhibited significant spectral changes. These observations were considered to be due to heterogeneity of lesions classified as LSIL and ASCUS.

Wood *et al.* (1996) investigated the use of principal component analysis (PCA) for data reduction and analysis. In a study of 272 patients, the observed spectra were visually sorted into two profiles: type 1 and type 2. Type 1 spectra exhibited a profile characteristic of normal epithelial cells with intense glycogen bands at 1022 and 1150 cm^{-1} and a pronounced phosphate stretch at 1078 cm^{-1}. Type 2 spectra exhibited features characteristic of malignant cells, with pronounced nucleic acids bands and a reduction in intensity of the glycogen bands, similar to those described by Wong *et. al.* (1993, 1991). Spectra of cultured malignant cells (from the HeLa cell line) exhibited a number of features common with type 2 spectra. PCA decomposition resulted in a score plot that showed general separation of the visually categorized spectra. 86% of the spectra that exhibited type 1 profiles by visual sorting and PCA analysis of F-IR spectra were diagnosed normal by Pap smear and 87% exhibiting type 2 spectral profiles were diagnosed abnormal.

Cohenford *et. al* (1997) also demonstrated the application of PCA to a study of cervical smears and suggested that blood may be an interfering substance. They found a sensitivity of 79% and specificity of 77%. Although specimens containing scraped exfoliated cells should not have much surrounding material except cells, as other studies came to light

(Wood *et al.* 1998, Wong *et al.* 2002) it became clear that other factors (such as the presence of blood in the sample) could be contributing to the spectral changes observed.

Fung *et al.* (1997) compared FTIR spectroscopy in the screening of cervical cells obtained through biopsy with cervical cells obtained through conventional Pap smears. Cells obtained from biopsy were found to be the gold standard. Exfoliated cervical cells from 301 patients were collected. The positive IR spectra were characterized by several or all changes in the spectra described by the previous studies outlined above. The sensitivity and specificity were 86.6% and 90.5% for Pap smears and 98.6% and 98.8% for FTIR spectroscopy. The Pap smear test showed 3.4% false negative and 9.5% false positive, compared to FTIR spectroscopy that showed 1.4% false negative and 1.2% false positive.

Shaw *et al.* (1999) showed that IR spectra of exfoliated cells carry information regarding the presence of dysplasia but concluded that automated feature extraction and pattern recognition methods are required to distinguish IR spectra of dysplastic smears from the spectra of normal smears. Chiriboga *et al.* (1998 a,b) demonstrated that FTIR spectroscopy could be used as a tool to monitor maturation and differentiation of human cervical cells and compared spectra of biopsies of squamous epithelia to exfoliated cervical cells. They further compared columnar tissue with squamous epithelium (Chiriboga *et. al* 1998 c) and found that features observed in the spectra of columnar cells in tissue samples were also detected in the spectra of exfoliated cells, indicating the presence of endocervical cells. It was shown that the columnar or glandular cells exhibit spectral features similar to those observed for pure cervical mucus. All four different layers in epithelial tissue and the stroma exhibit IR spectra distinguishable from each other. It was further concluded that

spectral features of healthy cervical cells are dominated by glycogen and proteins and that nucleic acids do not seem to contribute, especially in the case of fully mature cervical cells. A hypothesis was put forward that in a mature cervical cell the nucleus is opaque or nearly opaque in the mid-infrared range and that nucleic acid spectral features were due to RNA in the cytoplasm and not the DNA in the nucleus. This hypothesis was "verified" by an experiment in which an IR beam of ca. 5 μm diameter focused on a cell nucleus produced no IR light while a similar sized beam, focused on cytoplasm, transmitted a weak but an observable spectrum. It was also concluded, based on the spectra of single exfoliated cells that the glycogen concentration of single, mature cells can vary tremendously. Comparison of samples from different patients led to the conclusion that spectra are often contaminated by blood components (leukocytes and red blood cells), fungi or mucus.

Wood *et al.* (1998) reached a similar conclusion that bacteria, yeast, mucin, semen and other contaminants might contaminate FTIR spectra of smears.

Cohenford *et al.* (1998) examined individual cervical cells from normal, dysplastic and malignant samples. Cervical scrapings were processed using a Cytyc ThinPrep processor, which filters out mucus and debris and spreads the cells uniformly on a microscope slide. Ninety-five percent of normal exfoliated cells, comprised overwhelmingly of superficial and intermediate cells, displayed two distinct spectral patterns, designated A and B, while the remaining 5% displayed an intermediate pattern. Three main differences existed between the two patterns of spectra. In pattern B the amplitude of the band peaking around 1027 cm^{-1} was dramatically reduced compared with pattern A. The 1105 and 1153 cm^{-1} peaks were also shifted toward higher wavenumber; and the peak intensity ratio of the 1080

cm^{-1} band to the amide II band (composed mainly of NH in plane bending and CN stretching) was dramatically reduced. The following conclusions were reached: (1) superficial and intermediate squamous cervical cells exhibit no spectral difference; (2) both cell types display three distinct spectral patterns – A, B and an intermediate pattern; (3) spectra of cytologically normal cells from dysplastic or malignant cervical cells differ significantly from the spectra of cytologically normal cells from normal samples; and (4) parabasal, endocervical, koilocytic (cells with irregular hyperchromatic nuclei and fair cytoplasm), dysplastic and malignant cells show pattern B spectra. These findings suggested that the structural changes underlying the spectroscopic changes are involved in or are a product of cervical carcinogenesis and the neoplastic process may be more extensive than currently recognized with morphological criteria.

Neviliappan *et al.* (2001) studied the FTIR spectral features of normal and malignant exfoliated cervical cells, cells from malignant tissue and the cervical adenocarcinoma (CA) cell line (SiSo). Spectral bands of the CA tissue, exfoliated cells from CA and the cell line were all found to have similar spectral properties, but markedly different from that of exfoliated normal cervical cells. Significant changes in bands at $1025cm^{-1}$ (attributed to glycogen), $1080cm^{-1}$ (glycogen and nucleic acids), $1155cm^{-1}$ (C-OH groups of serine, threonine and tyrosine of cell proteins and C-O groups of carbohydrates), $1240cm^{-1}$ (PO_2^- groups of nucleic acids), $1400cm^{-1}$ (methyl group of lipids and proteins) and $1450cm^{-1}$ (methylene groups of lipids and proteins) were noted in the CA tissue, exfoliated CA cells and adenocarcinoma cell line compared with exfoliated normal cells. Marked shifts in band position from 1080 to $1086cm^{-1}$, 1153 to $1160cm^{-1}$ and 935 to $970cm^{-1}$ were observed in CA tissue, exfoliated CA cells and the adenocarcinoma cell line. Overall the spectral bands

of the adenocarcinoma cell line were found to correlate well with those of cervical CA tissue and exfoliated CA cells, but to differ greatly from those of normal exfoliated cells.

Wong *et al.* (2002) examined the confounding variables that can lead to a misinterpretation of the FTIR spectra of exfoliated cervical cells. The study details the effect of variables such as; polymorphs (white blood cells), cell degradation and impurities such as cervical columnar cells, metaplastic cells (immature squamous cells), cervical mucus, red blood cells and debris. For example, apart from obscuring cytopathology, the presence of polymorphs in a cervical cell specimen interferes with the FTIR analysis of the spectrum. In relation to sample degradation, the study found that cervical specimens stored in saline solution at room temperature for a number of hours lost their original cellular structure and also resulted in changes in the FTIR spectrum (such as broadening of the glycogen band and the band at $1080cm^{-1}$ appearing as a shoulder). The presence of columnar cells (10%) amongst normal ectocervical squamous cells was also examined and was found to cause only insignificant changes. The effect of the same quantity of metaplastic cells was also found to be negligible. However the presence of cervical mucus (20%) in the sample was found to result in a decrease in the relative intensity of the glycogen band at $1023cm^{-1}$ versus the $1082cm^{-1}$ band. This suggests that if only the glycogen band region of the IR spectra is used as a spectroscopic screening method, small amounts of mucus could lead to an incorrect interpretation of the FTIR spectrum.

Romeo *et al.* (2002a) carried out a study to examine the IR spectral properties of a number of different cervical cell types obtained from cervical smears. Cells examined included, normal, abnormal (cervical intraepithelial neoplasia - CIN II and III), carcinoma *in situ*,

normal endocervical, inflammation, *Candida albicans* (yeast infection) and bacterial vaginosis. Spectral differences in all cell and diagnostic types investigated were found in the phosphodiester and carbohydrate regions (1000-1250cm^{-1}). However the spectral differences in other bands were not distinct enough to allow differentiation between groups. Differences between the endocervical and ectocervical spectra were attributed to increased mucin levels in endocervical columnar cells. Spectra of normal ectocervical cells compared to abnormal ectocervical cells and normal endocervical cells displayed increased lipid, protein and glycogen band intensity. Yeast infection was characterised by an intense broad band in the polysaccharide region (900-1150cm^{-1}) and pronounced $\nu_{as}PO_2^-$ and δCH_3 bands, although these properties were not exhibited by the averaged samples in the study and it was therefore deemed to be a function of the severity of the infection. Cells with bacterial vaginosis (infection without inflammation), displayed spectral differences in the phosphodiester and carbohydrate regions. PCA was successfully used to obtain a separation of normal ectocervical smears from normal endocervical smears and smears diagnosed with inflammation, *Candida* and bacterial vaginosis. A separation, with a slight overlap, of abnormal ectocervical smears from normal endocervical cells, inflammation and bacterial vaginosis was also obtained with PCA. *Candida* was not separated from abnormal ectocervical smears with any success.

The study by Diem *et al.* (2002) examined the IR spectra of individual normal and cancerous cells with a spatial resolution that permitted distinction between the nuclear and cytoplasmic regions. The study highlights the fact that the diagnostically significant peak at 1050cm^{-1} present in both immature and neoplastic cells, is obscured by the consistently

high glycogen content seen in both superficial and intermediate squamous cells. The study also warns that actively dividing/metabolically active cells have spectra that exhibit distinct DNA/RNA contributions seen between 1000-1250cm^{-1} (a possible indication of a diseased state). However cells with similar spectra were seen in both healthy and diseased patients, therefore this observation cannot alone be used as an indication of cancerous cells. It was found however that the number of cells with spectral features of immature cells is distinctly higher in samples of women with dysplasia and among these cells, nucleic acid features were stronger. The study suggests that a reliable detection of cervical dysplasia is possible if spectra are collected on an individual basis.

A short review of the area by Mantsch *et al.* (2002) posed the question; "vibrational spectroscopy and medicine; an alliance in the making?" The paper discussed many different aspects of the field. In relation to cervical cancer, the authors warn physical scientists working in the area not to underestimate the complexity of the biological systems under examination. Many of the inherent difficulties in relation to cervical cancer diagnosis are outlined such as the presence of contaminants including blood and mucus, the fact that epithelial cells themselves may arise from different locations within the epithelial layer and the fact that the cervical cells in a smear sample are heterogeneous by their nature.

Chang *et al.* (2003) examined 22 frozen cervical biopsy specimens in a study to detect cervical precancerous tissue using FTIR spectroscopy. The study found that the region from 950-1500cm^{-1} was the most critical region in the IR spectrum of tissue. Peaks identified as diagnostically useful included those at 1150 and 1240cm^{-1} which were found to decrease and broaden respectively with the onset of pathological changes. Also, the ratio of the areas

at 1130-1180cm^{-1} and 1180-1260cm^{-1} was found to be a useful factor in discriminating precancerous tissues from normal tissues.

Wood *et al.* (2003) conducted a study examining FTIR mapping of the cervical transformation zone and dysplastic squamous epithelium. Photomicrographs were taken of the region under investigation and 90 spectral maps were recorded from various regions of interest. Unsupervised hierarchical cluster analysis was performed on the spectra and the resulting pseudo-colour spectral map was compared to anatomical features seen in parallel haematoxylin and eosin (H+E) stained sections. Several distinct regions in the pseudo-colour map could be directly correlated with the main tissue layers in the H+E stained section. These included the outer superficial layer, the intermediate layer, the parabasal and basal layers and the connective tissue. Strong glycogen contributions were seen in the squamous epithelial layers, while the basal/parabasal and stroma layers were differentiated by their protein bands. Cervical intraepithelial tissue (CIN) was also examined and it was found to have several characteristic features including pronounced symmetric and asymmetric phosphate stretching modes at 1078 and 1240cm^{-1}, a significant reduction in glycogen band intensity and a relatively small Amide II / Amide I ratio.

Mordechai *et al.* (2004) examined both cervical and skin tissues for common biomarkers of cancer. The study found carbohydrate levels (glycogen) showed a good diagnostic potential for the detection of cervical cancer but not for melanoma. However increases in the nucleic acid concentration in both tissue types (measured by calculating the ratio of the intensities of the bands at 1121 and 1020 cm^{-1} i.e. I_{1121}/I_{1020}) was found to be a possible common biomarker.

Romeo *et al.* (2004) investigated the FTIR spectra of individual cultured human cervical cells (HeLa cell line). Spectra were collected microscopically in reflection/absorption modes, from cells deposited and dried on microscope slides or from cells grown directly on slides. The study warns that an understanding of cell-growth-related effects is required for an eventual cytological application of IR spectroscopy. Within the spectra of dried cells, significant heterogeneity was seen. This was mainly composed of changes in Amide I / Amide II band intensity ratios and Amide I to phosphate stretching band intensity ratios (1080-1240cm^{-1}), as well as slight shifts in frequencies. The paper postulates that the changes in intensity and frequency shifts in the spectra of dried cells may not be entirely due to different progression points of cells through the cell division cycle, as spectra of cells of similar sizes display a significant overall intensity variation. The study suggested that these intensity changes must be due to differences in the concentration of biochemical components and this in turn may be influenced by the morphology of the cells as they dried.

Diem *et al.* (2004) conducted a review of a decade of vibrational micro-spectroscopy of human cells and tissues. The paper charts the emergence of the field from the early work dating back to the late 1980's and early 1990's. Early studies encountered problems with macroscopic examination of cells and tissues, where homogeneity of the sample was not assured. These problems have been overcome by using microspectroscopic methods along with a high quality visual image of the exact spot from where the spectra are recorded. A variety of different tissue/cell types, as well as two-dimensional spectral maps were examined by the paper. The spectral characteristics of cervical squamous intermediate epithelium, squamous basal layer and columnar epithelium are displayed. The spectral

features of glycogen were clearly present in the spectra of squamous epithelial cells. The study discussed the presence of glycogen in these cells, with bands seen at 1151, 1078 and 1028cm^{-1} attributed to coupled C-O and C-C stretching and C-O-H deformation motions. Variations in the glycogen content as a method of utilising IR microspectroscopy as a screening tool was discussed. Diem *et al.* warned that glycogen depletion, although seen in dysplastic or cancerous tissue, is non-specific and many normal, superficial squamous cells in the final stage of their life cycle completely deplete their glycogen storage. Rapidly proliferating cells also deplete their glycogen. Thus, glycogen content alone should not be used as an indicator for disease.

With this large body of research having been carried out into the area of cervical cancer and FTIR spectroscopy, the present study aims to reproduce some of the results seen previously, as well as identifying any additional markers of disease. The study will be used as a comparative study with Raman spectroscopy. This present study will employ a new method of tissue dewaxing (as outlined in the previous chapter) and will utilise synchrotron FTIR spectroscopy, with its greater spatial resolution than conventional FTIR spectroscopy. Conventional FTIR spectrometers use blackbody sources, such as a tungsten-halide lamp to generate the excitation radiation. Synchrotron electromagnetic radiation, on the other hand, is emitted tangentially by electrons moving at relativistic speed in a circular storage ring and accelerated by means of bending magnets (Fig 5.1). The advantages associated with the use of synchrotron radiation for infrared (SR-FTIR) spectroscopic measurements include the provision of a small effective source size and low thermal noise (due to the low-divergence beam from the IR synchrotron source). This results in a stable, highly collimated beam produced at the experimental station (Bozec *et al.* 2002).

The recent development of FTIR capabilities at synchrotron laboratories has facilitated the study of biological samples about 5 – 20 μm in dimension (Tobin *et. al* 2002), with a beam up to 100 times more intense than is possible using a conventional IR source (Miller *et. al* 1997). These advantages, coupled with a spatial resolution closer to that of Raman spectroscopy than conventional bench-top FTIR, make SR-FTIR suitable for a comparative study such as this.

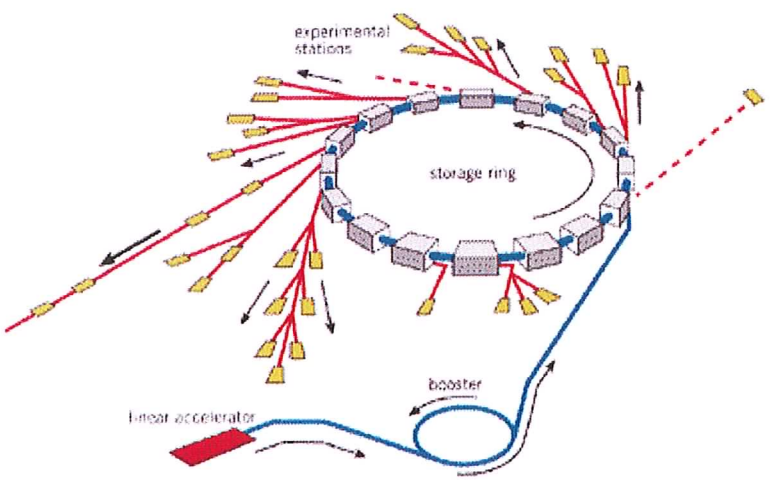


Figure 5.1 Schematic of synchrotron radiation source, including the many different stations delivering radiation with wavelengths extending from the infrared to X-rays

(Daresbury Synchrotron Radiation Source)

The aim of the study is to examine the SR-FTIR spectra of dewaxed FFPP cervical tissue sections and to determine whether different cell types exhibit sufficient spectral specificity to allow differentiation based on their spectral characteristics. Different cell types seen in normal cervical sections were recorded, as well as spectra from invasive carcinoma tissue

and premalignant neoplastic tissue. Differences between cell types, as well as differences between normal and abnormal tissue are discussed.

5.2 Materials and methods

5.2.1 Instrumentation

Synchrotron FTIR measurements were taken at the 13.3 beamline at the synchrotron radiation source (SRS), Daresbury Laboratory as outlined in Chapter 1.3.2. Various parameters were investigated to obtain a satisfactory spectrum. The microscope aperture was set to 10μm × 10μm and 128 scans with an $8cm^{-1}$ resolution yielded good quality spectra.

Due to the fact that glass absorbs in the IR region, regular glass slides were not suitable for mounting tissue sections. IR reflective slides (MirrIR) were obtained from Kevley Technologies (Chesterland, Ohio, USA). Low-e MirrIR microscope slides are glass microscope slides that are coated with a thin Ag/SnO_2 layer. They are chemically inert and nearly transparent to visible light. However, they reflect near infrared radiation nearly completely. Thus, they are an ideal (and inexpensive) substrate for reflection/absorption infrared microspectroscopy, since visual images and IR spectra can be obtained from the same sample (Romeo *et al.* 2004). The MirrIR slides were evaluated for deterioration resulting from exposure to chemicals used in the dewaxing process and minimal deterioration in its IR reflective properties was found.

5.2.2 Sample preparation (FFPP sections)

40 formalin fixed paraffin processed (FFPP) cervical tissue sections were obtained from the National Maternity Hospital (NMH), Holles Street, Dublin. As characterised by the Registrar, the samples received consisted of 20 normal and 20 invasive carcinoma. Of the 20 carcinoma sections received, 10 sections were identified as having various grades of cervical intraepithelial neoplasia (CIN), which were also marked for examination. After the wax impregnation, tissue was embedded and sliced into 10μm sections, mounted on MirrIR slides (Kevley Technologies, Chesterland, Ohio, USA) and dried. Two sections were taken from each sample. The sections were immersed in a series of baths consisting of a bath of hexane (BDH, Dorset, UK) overnight, two baths of ethanol absolute (Merck, Dorset, UK) for three minutes and two minutes and a final bath of industrial methylated spirits 95% (Lennox, Dublin, Ireland) for one minute. One section from each sample (the reference section) was stained with Haematoxylin and Eosin and the other kept unstained for spectroscopic examination.

5.3 Results

5.3.1 FFPP Normal Cervical Tissue

Due to the fact that residual wax contributions from dewaxed FFPP sections proved problematic during Raman spectroscopic investigations, the FTIR spectral properties of wax were initially examined. Figure 5.2 shows the absorption spectrum of paraffin wax and

the infrared spectrum of an ineffectively dewaxed FFPP section (taken prior to the use of hexane). The contributions from the wax are visible in Figure 5.2(B)

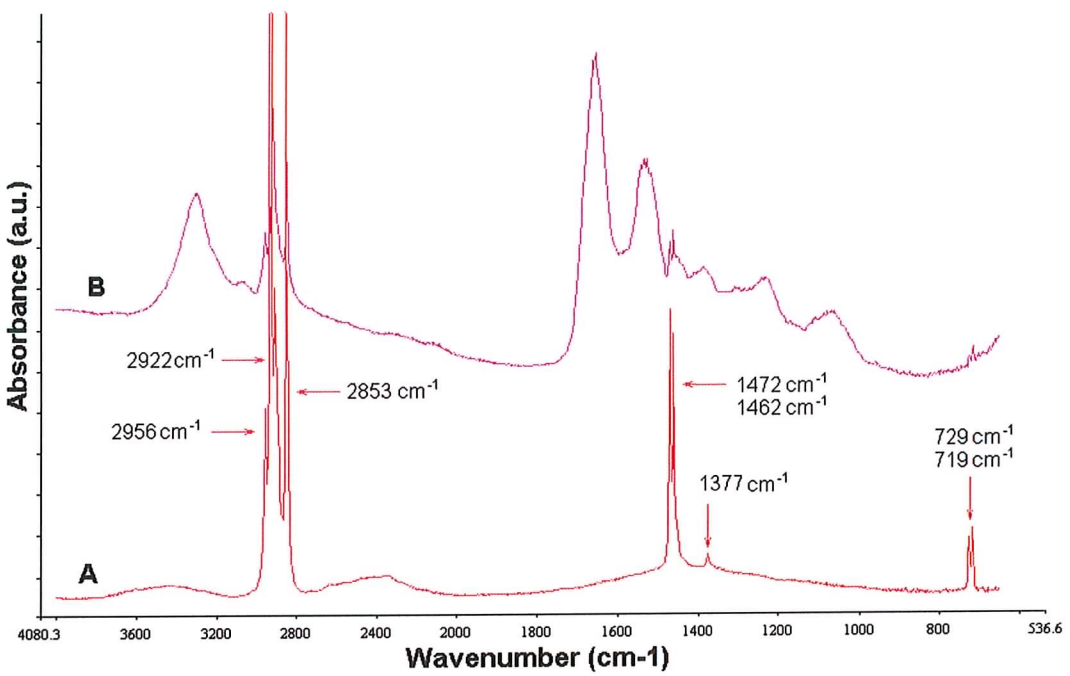


Figure 5.2 FTIR absorption spectra of A) paraffin and B) dewaxed FFPP section

Contributions were observed at 2956 cm^{-1}, 2922 cm^{-1} and 2853 cm^{-1} which correspond to CH_3 asymmetric stretch, CH_2 asymmetric stretch and CH_2 symmetric stretching modes respectively. Peaks at 1462 cm^{-1} and 1472 cm^{-1} correspond to CH_2 scissoring modes. The small band at 1377 cm^{-1} is the C=O symmetric stretch of COO^- and the two peaks at 719 cm^{-1} and 729 cm^{-1} correspond to C-H rocking.

Although many spectra were recorded only a representative selection are shown. Figure 5.3 shows the spectra of three different cell types of dewaxed FFPP tissue sections from normal cervix. Spectra have been offset to facilitate clarity and are shown over the entire spectral

range (400 – 4000 cm^{-1}). As seen in Raman spectroscopy, contributions at frequencies greater than ≈ 1800 cm^{-1} are mainly due to CH_2 and CH_3 stretching vibrations and are greatly affected by residual wax contributions. Also there are no contributions of interest below ≈ 800 cm^{-1}. The fingerprint region ranges from ≈ 800 cm^{-1} - 1800 cm^{-1} and it was this region that was closely examined.

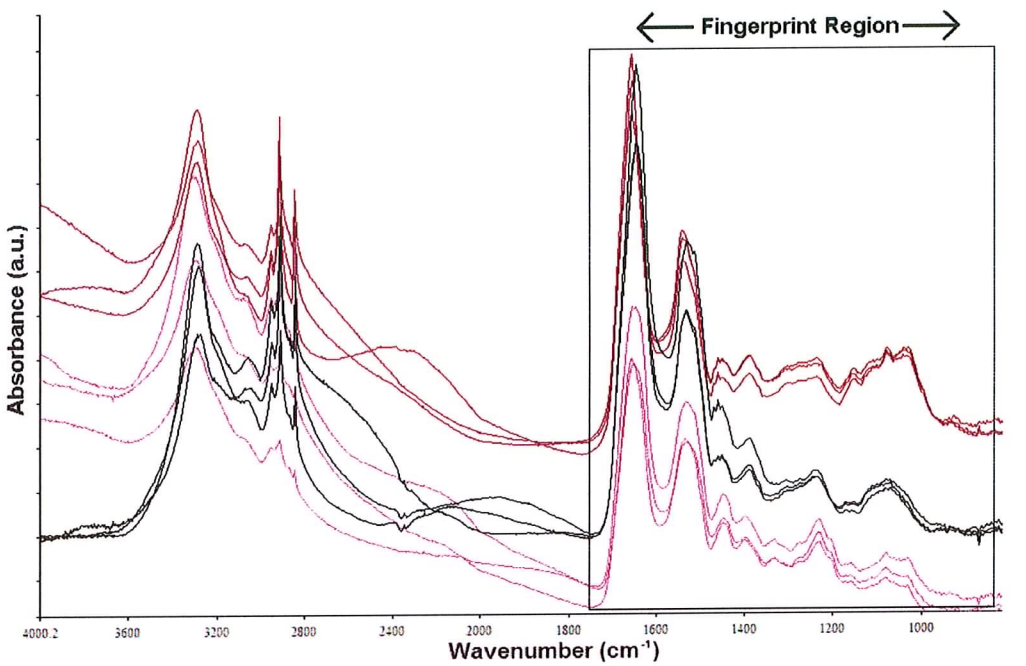


Figure 5.3 FTIR spectra of FFPP sections over the entire spectral range (400 – 4000 cm^{-1})

A table outlining FTIR assignments from a biological perspective is given in Table 5.1

Approximate Frequency (cm^{-1})	Assignment
3500	O-H str of hydroxyl groups
3200	N-H str (Amide A) of proteins
2959	C-H str (asym) of $-CH_3$ methyl
2934	C-H str (asym) of $>CH_2$ methylene
2921	C-H str (asym) of $>CH_2$methylene in fatty acids
2898	C-H str of methine
2872	C-H str (sym) of $-CH_3$ methyl
2852	C-H str (asym) of $>CH_2$ methylene in fatty acids
1741	>C=O str of esters
1715	>C=O str of ester, protonated carboxyl and carbonyl groups
1695	different amide I band components resulting
1685	from antiparallel pleated sheets
1675	and β-turns of proteins
1655	amide I of α-helical structures
1637	amide 1 of β-pleated sheet structures
1548	amide II band
1515	tyrosine band
1468	C-H sc of $>CH_2$ methylene
1400	C=O symmetric stretch of COO^-
1310-1240	amide III band components of proteins
1250-1220	P=O str (asym) of $>PO_2^-$ phosphodiesters
1088-1084	P=O str (sym) of $>PO_2^-$ phosphodiesters
1200-900	C-O-C, C-O dominated by the ring vibrations of carbohydrates,C-O-P, P-O-P
720	C-H rocking of $>CH_2$ methylene
sc = scissoring; str = stretching; sym = symmetrical; asym = asymmetrical	

Table 5.1 Assignment of some of the spectral bands frequently found in FTIR biological specimens (assignments from Naumann 2001)

Figure 5.4 examines the fingerprint region of Figure 5.3 more closely. Spectra were recorded from a single sample of normal cervix. The figure shows the infrared spectra of three classes of tissue, namely basal layer, connective tissue and epithelial tissue. The three tissue classes have been offset to improve clarity. A micrograph of the connective tissue, basal and epithelial cells is shown in Figure 5.5. The dominant peaks have been labelled on the graph for each of the three tissue types. The 3 different tissue types exhibit different spectral properties. The most obvious distinction between the different classes of tissue are the features exhibited by the epithelial cells in the 1100 cm^{-1} region. Three sharp peaks at 1036 cm^{-1}, 1079 cm^{-1} and 1152 cm^{-1} and a further peak at 930 cm^{-1} were present in normal epithelial cells.

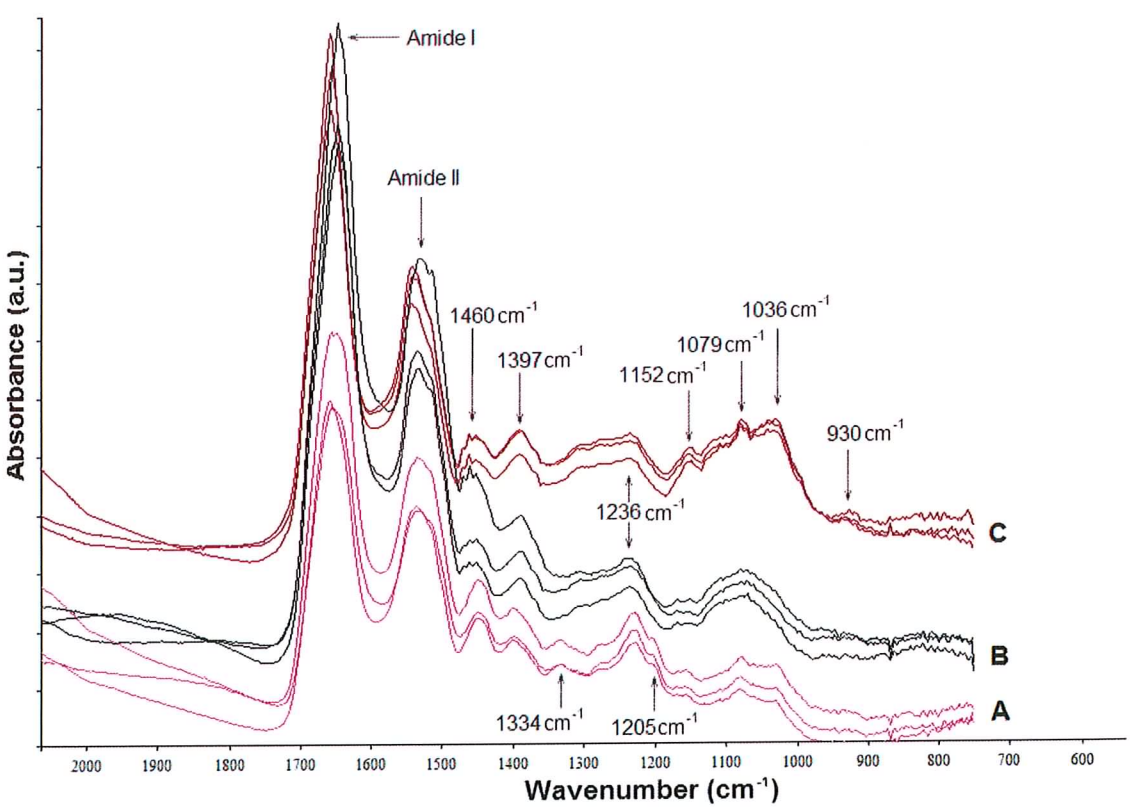


Figure 5.4 Absorption spectra of A) connective tissue B) basal cells and C) epithelial cells

The band at 1079 cm^{-1} originates from the symmetric stretch of phosphodiester groups of nucleic acids (Cohenford *et al.* 1997). The bands at 1036 cm^{-1} and 1152 cm^{-1} are attributable to glycogen (Wong *et al.* 1991). They are derived mainly from vibrational modes of CH_2OH groups of carbohydrates (see Chapter 2, Fig 2.12). Other bands seen include; C-H scissoring of $>CH_2$ (1460 cm^{-1}), C=O symmetric stretch of COO^- (1397 cm^{-1}) and P=O asymmetric stretching of $>PO_2^-$ phosphodiesters (1236 cm^{-1}) and were present in all three classes of tissue (Naumann 2001). The positioning of the Amide I band, relating to protein conformation was found to vary over a 20cm^{-1} range between all the cell types.

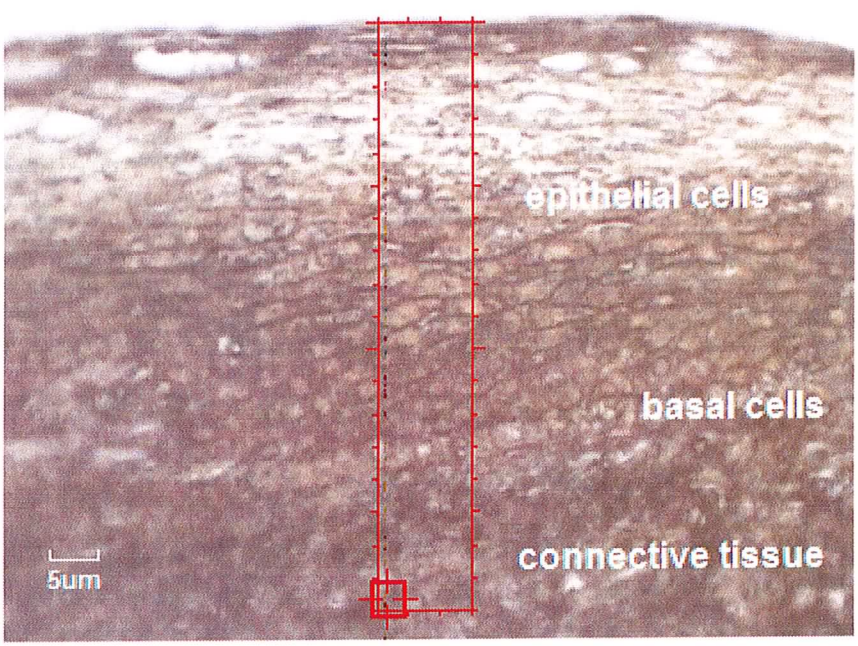


Figure 5.5 Photomicrograph showing connective tissue, basal and epithelial cells

Figure 5.6 examines the spectral properties of the different types of epithelial cells, (parabasal, intermediate and superficial squamous). Again, the spectra have been offset.

Increases in intensity of the bands at 1036 cm^{-1}, 1079 cm^{-1} and 1152 cm^{-1} were seen as the cell matured. The Amide I band was seen to vary in intensity and shift over a ≈10cm^{-1} range. The band at 930cm^{-1} is again seen in the epithelial cells and the band was seen to increase in intensity from parabasal through to superficial squamous cells. A micrograph is also shown in Figure 5.7.

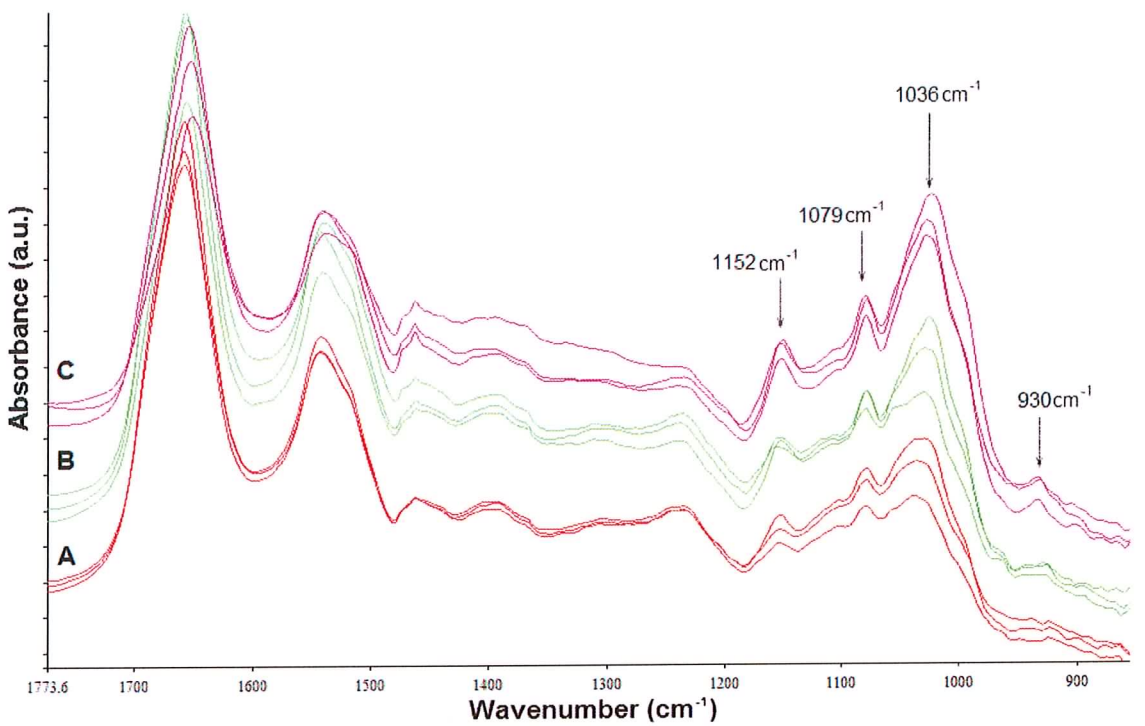


Figure 5.6 FTIR spectra of A) parabasal, B) intermediate and C) superficial epithelial cells

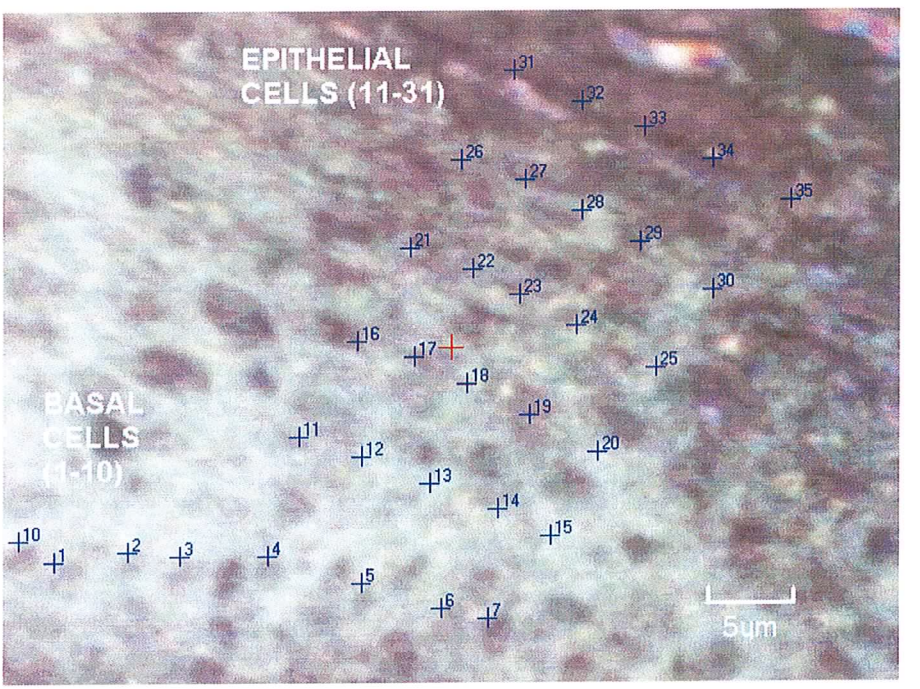


Figure 5.7 Photomicrograph of cells examined using x32 objective lens

Figure 5.8 shows the IR spectra from a single patient recorded from 10 regions of intermediate epithelial cells to examine spot to spot variation. The spectra display a good degree of consistency, with the exception of the fluctuation of the different components that make up the Amide I band as seen previously. The bands at 1152, 1079, 1036 and 930 cm^{-1} were present in all spots sampled.

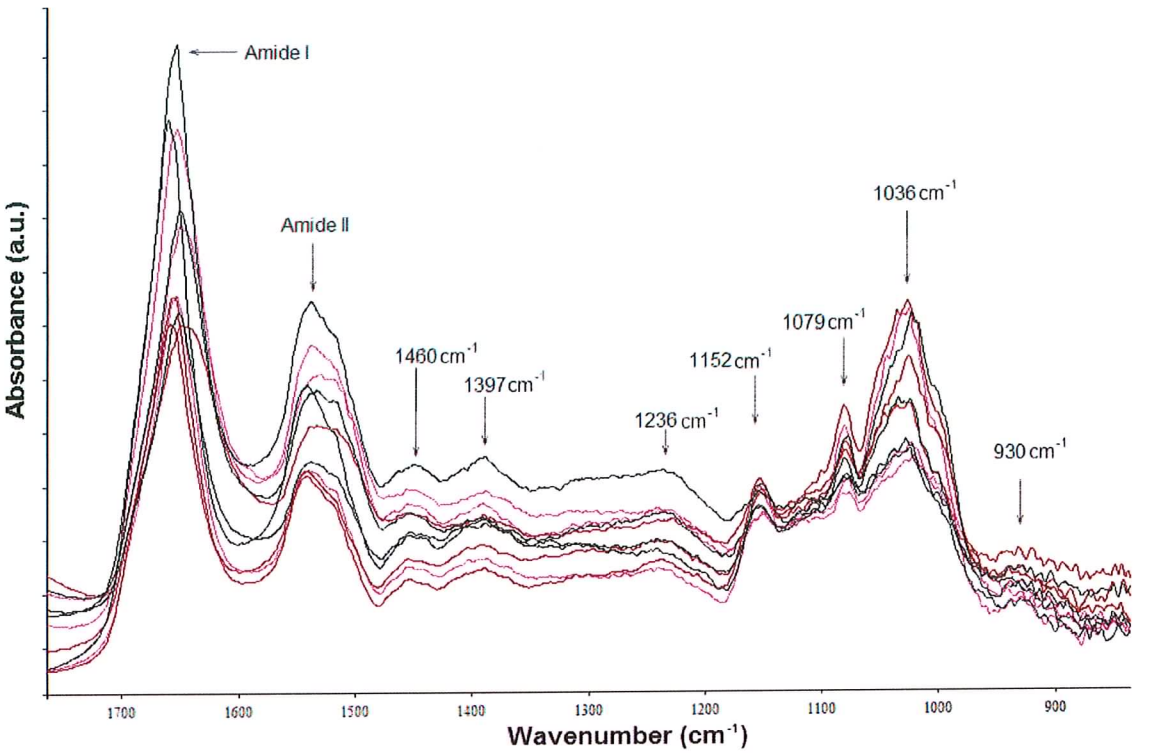


Figure 5.8 IR spectra of a single patient recorded from intermediate epithelial cells

Figure 5.9 examines the interpatient variability between 10 normal tissue samples. Although different samples displayed changes in overall intensity, strong correlation was seen between the samples at 1036, 1079, 1152 and 930 cm^{-1}. Interpatient samples resulted in large variations in the composition of the Amide I band and it is clear that a number of samples have a small amount of residual wax, although a large improvement was observed compared to previous samples dewaxed using xylene.

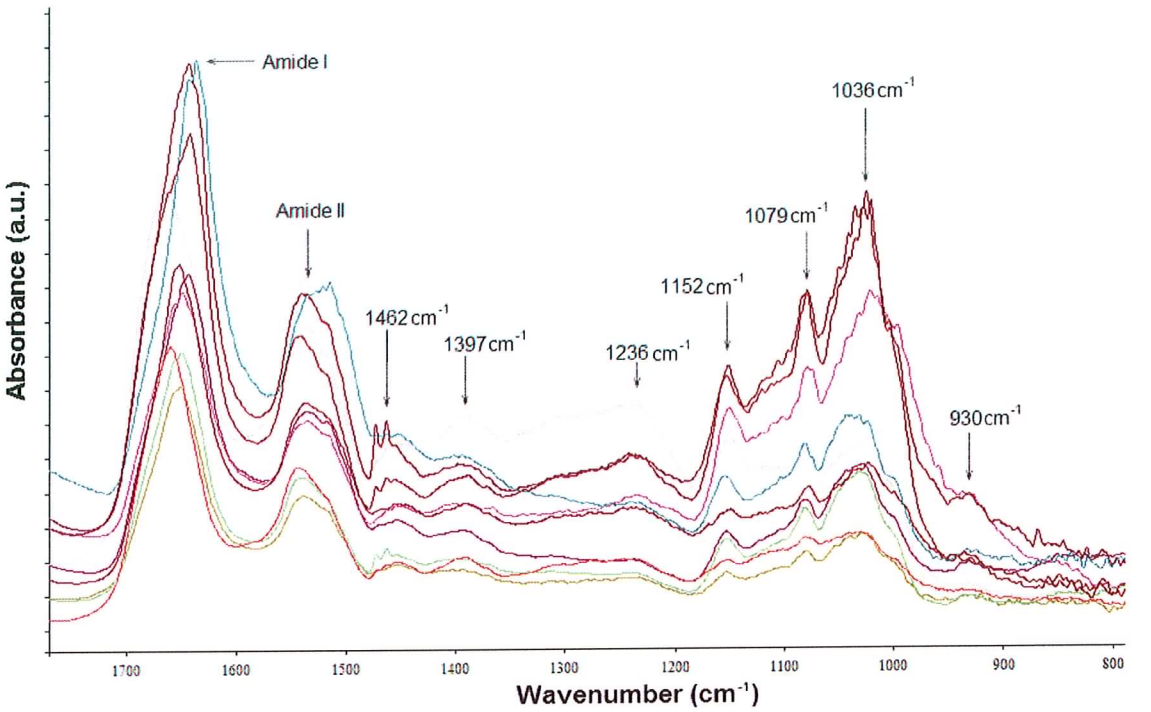


Figure 5.9 IR spectra from 10 patients recorded from intermediate epithelial cells

5.3.2 FFPP Invasive Carcinoma

Figure 5.10 is a photomicrograph showing 10 spots recorded from a single sample of invasive carcinoma. The resulting IR spectra are shown in Figure 5.11. The spectra of the invasive carcinoma tissue display a marked difference to the spectra recorded from normal epithelial tissue. The most noticeable differences include a stronger contribution at 1236 cm^{-1} and a loss of the 1036, 1079, 1152 cm^{-1} triad, which is replaced by a single, weak peak at 1079 cm^{-1}. Also the peak seen at 930 cm^{-1} is not evident and is instead replaced by a peak at 970 cm^{-1}. Amide I and II bands are present, although the Amide I band shifts at different

spots is markedly less than seen with normal tissue. Other notable peaks include those at 1462 (C-H scissoring of $>CH_2$), 1397 (C=O symmetric stretch of COO^-) and 1169 cm^{-1} (C-O stretch).

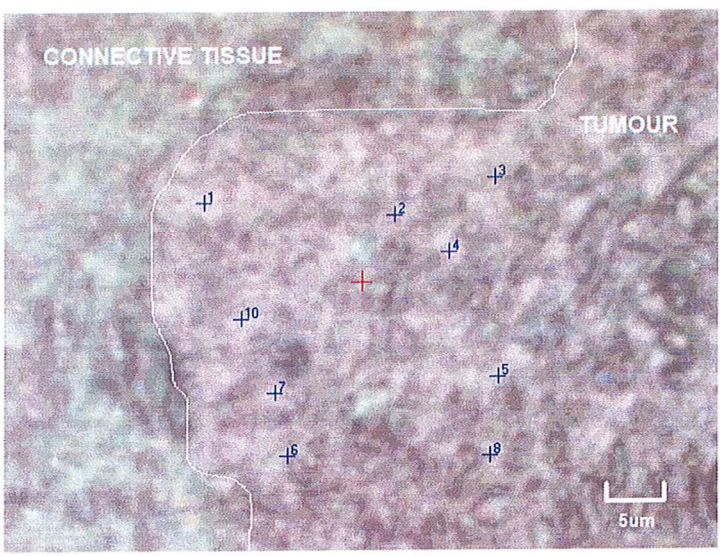


Figure 5.10 Micrograph showing 10 spots recorded from invasive carcinoma

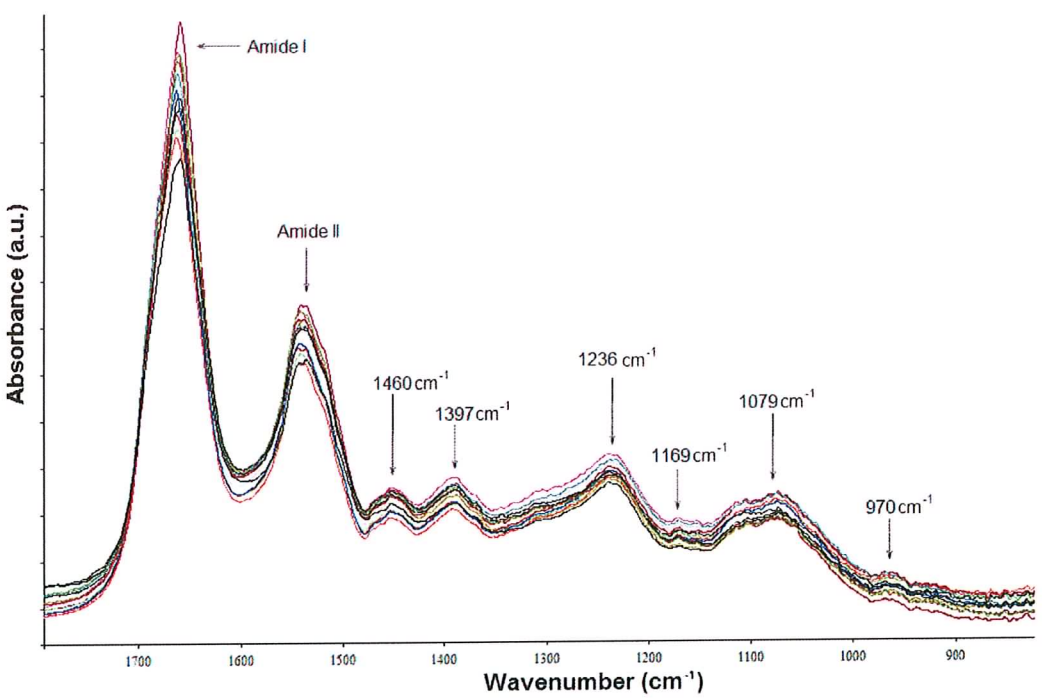


Figure 5.11 IR spectra recorded from a single patient to examine spot to spot variation in carcinoma samples.

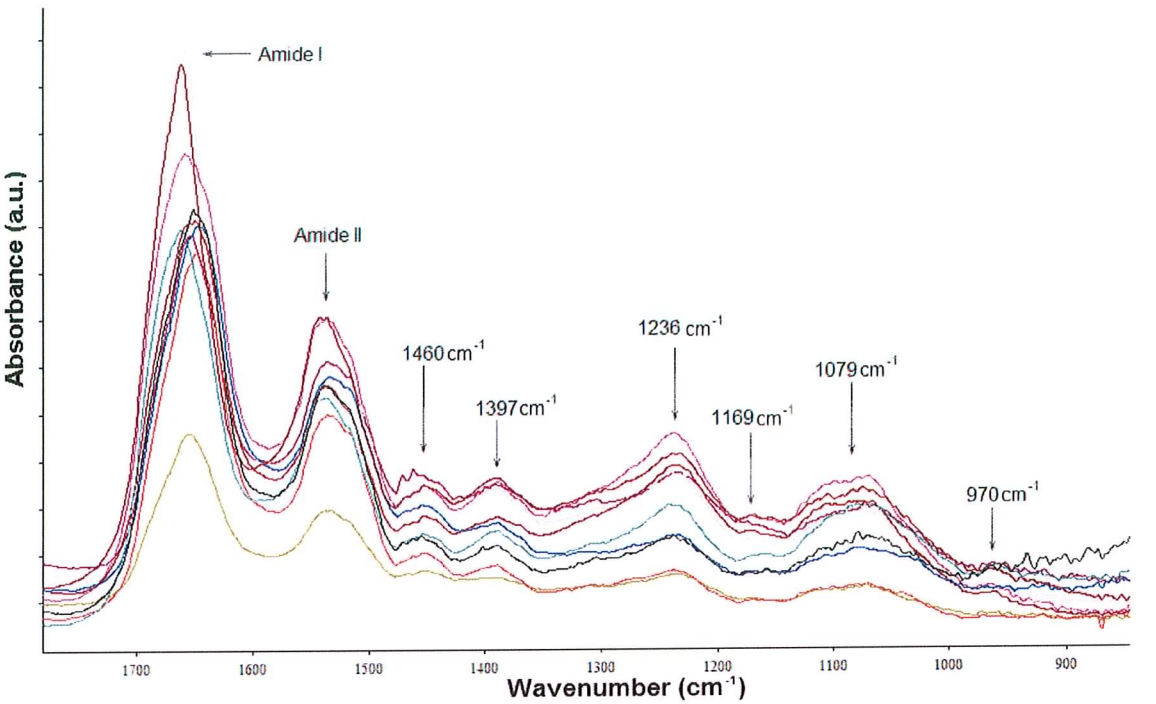


Figure 5.12 IR spectra recorded from 10 different patients to examine interpatient variability

Spot to spot variation was found to be minimal. Figure 5.12 shows 10 IR spectra recorded from 10 different carcinoma samples. Although the spectra recorded from a single sample of invasive carcinoma showed good correlation of the Amide I band, it can be seen from Figure 5.12 that this does not apply to the interpatient samples examined, where large variations in the positioning of the Amide I band were again apparent. All the samples exhibited similar spectra, with peaks seen at 1460, 1397, 1236, 1169, 1079 and 970 cm^{-1}.

5.3.3 FFPP Neoplastic Tissue

Figures 5.13 and 5.14 respectively show a micrograph of points sampled and the infrared spectra of a section diagnosed as severe cervical intraepithelial neoplasia (CIN). The micrograph shows undifferentiated neoplastic cells which have taken over the normal epithelial cells in the epithelial layer. This neoplastic growth is particularly obvious when compared with the growth seen in normal epithelium in Figure 5.5. The spectra recorded from this area show considerable similarity to those recorded from areas of invasive carcinoma. Spot to spot variation was found to be minimal. Figure 5.15 examines the variation between different samples and the degree of interpatient variability is minimal, with the exception of the Amide I band which shifts over a $10cm^{-1}$ range.

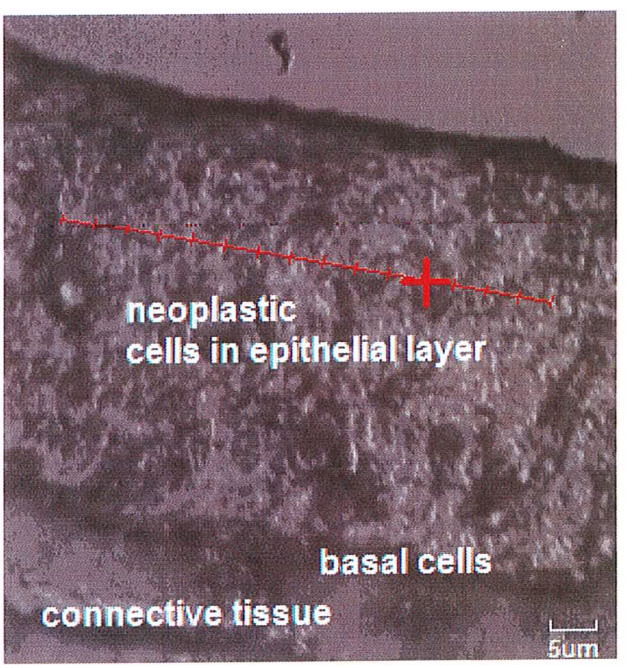


Figure 5.13 Micrograph showing CIN region

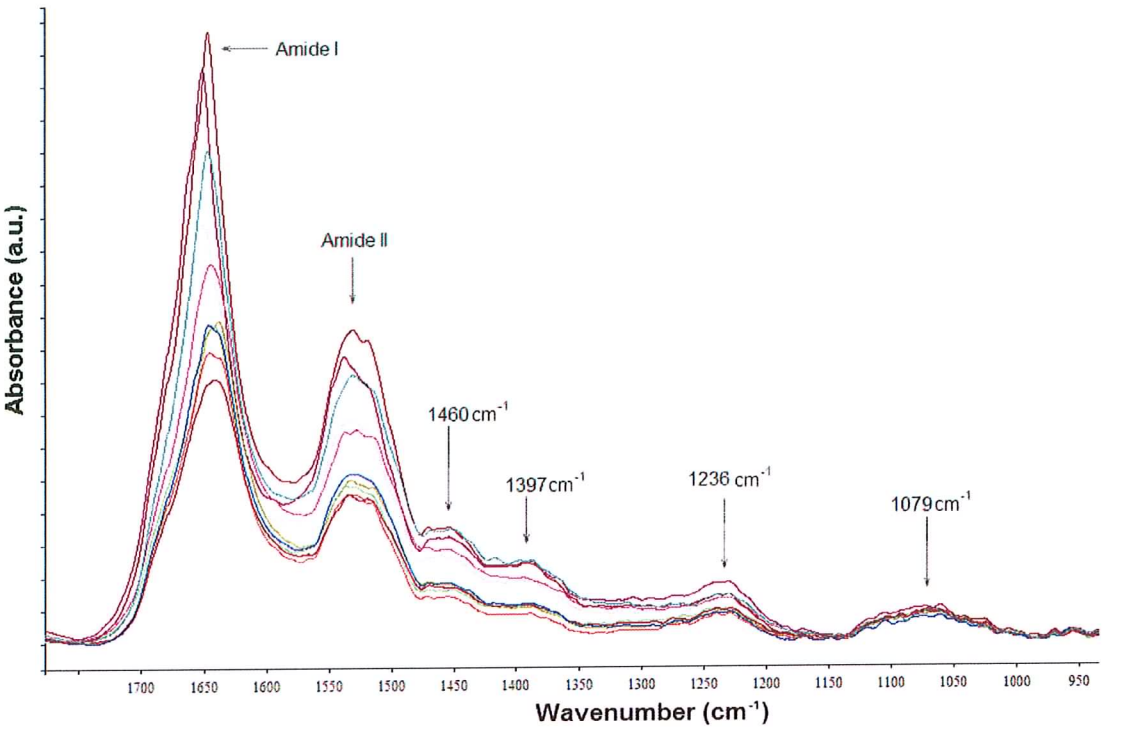


Figure 5.14 IR spectra recorded from 10 different regions of a single CIN sample

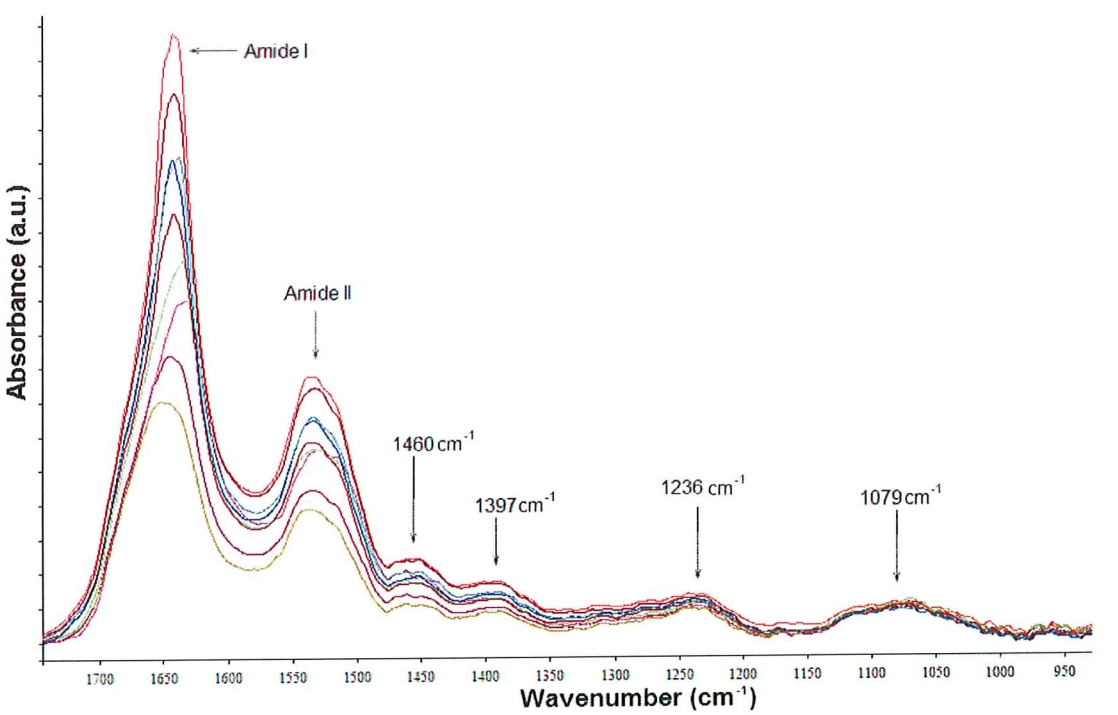


Figure 5.15 IR spectra recorded from 10 CIN samples to examine interpatient variability

Figure 5.16 directly compares the infrared spectra of invasive carcinoma and normal epithelial cells. In both samples the Amide I band is disorderly and there is a shift of ≈ 10 cm^{-1} between the Amide II band of the epithelial and cancer cells. All other differences with carcinogenesis were replicated, including;

a dramatic decrease in the band associated with glycogen at 1036 cm^{-1}

a decrease in the narrow glycogen band at 1079cm^{-1}, replaced with a less intense, broader peak

a decrease in the band associated with glycogen at 1152 cm^{-1}

a small increase in the nucleic acid band at 1236 cm^{-1}

a broadening of the peak associated with CH_2 at 1462 cm^{-1}

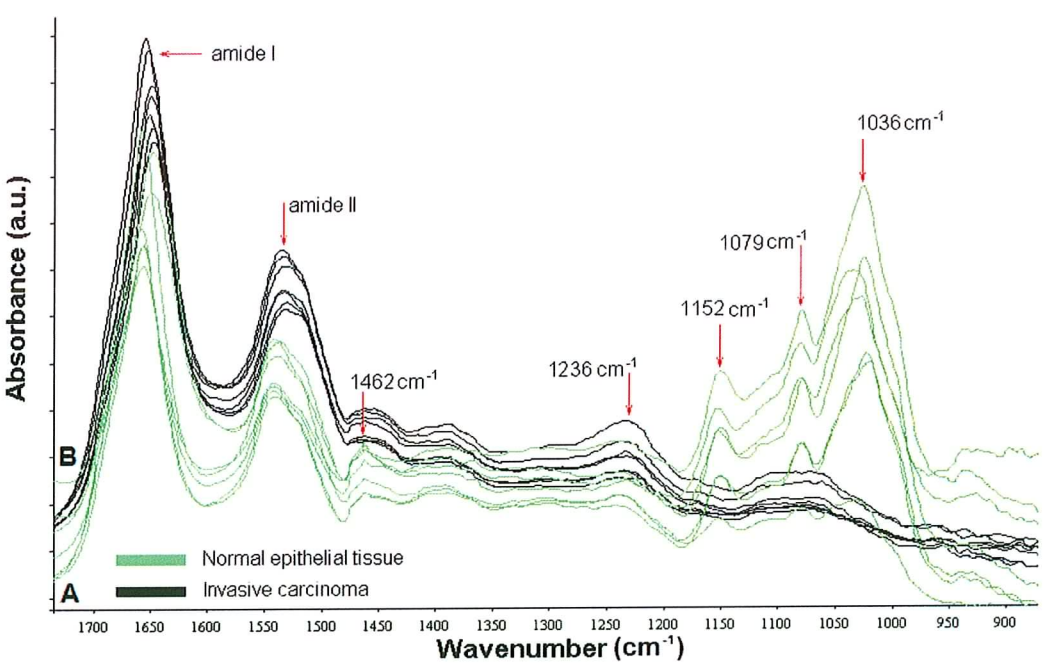


Figure 5.16 IR spectra of A) normal epithelial tissue and B) invasive carcinoma

Over 400 spectra were recorded from the 40 dewaxed FFPP cervical sections. However for discussion purposes only a representative selection have been shown here. All of the samples and cell types examined displayed broadly similar infrared spectral characteristics as outlined above.

5.4 Conclusion

As with Raman spectroscopy, it was possible in the IR study to identify the peaks resulting from the presence of wax. For the purposes of the present study, the most important peaks were seen in the fingerprint region at 1462 cm^{-1}, 1472 cm^{-1} and 1377 cm^{-1} (CH_2 scissoring and C=O symmetric stretch of COO^- respectively) and should not be used in a diagnostic capacity.

The spectral properties of connective tissue, basal cells and epithelial cells were examined. SR-FTIR was found to provide many diagnostically useful peaks relating to the composition of the biological material under investigation, although features seen are much broader than those seen using Raman spectroscopy. However it is clear that each cell type displayed its own distinct spectrum. Peaks identified common to all cell types included 1660 and 1545cm^{-1} (Amide I and II bands), 1460cm^{-1} (C-H scissoring of CH_2), 1397cm^{-1} (C=O symmetric stretch), 1236cm^{-1} (P=O asymmetric stretch) and 1079cm^{-1} (P=O symmetric stretch).

The Amide I band of proteins, which mainly involves the carbonyl stretching vibrations of the peptide backbone, is a sensitive marker of protein secondary structure, as the vibrational frequency of each C=O bond depends on hydrogen bonding and the interactions between the Amide units, both of which are influenced by the secondary structure. The Amide I band was found to shift over a range of $\approx 10 cm^{-1}$ both within each class and between each class of tissue indicating changes in the secondary structure.

The Amide II band is also affected by the different elements of secondary structure. However it is not as sensitive a marker to secondary structure content. This is due to the fact that the Amide II mode is a heavily mixed vibrational mode. The Amide II mode is an out-of-phase combination of NH in-plane-bending and CN stretching with minor contributions from C=O in plane bending, CC stretching and NC stretching (Arrondo *et al.* 1993). In the present study the position of the Amide II band was found to be much more consistent than that of the Amide I band.

In addition to the peaks outlined above, spectra of connective tissue displayed peaks such as the peak seen at $1036 cm^{-1}$ (CH_2OH stretch). Two other characteristic peaks were seen at 1334 and 1205 cm^{-1}. These peaks are attributable to collagen (Wong *et al.* 1995).

Basal cells gave spectra which showed strong and broad bands at 1236 and $1079 cm^{-1}$. These bands are due to asymmetric and symmetric stretching of phosphate (PO_2^-) modes. These vibrations originate mainly in the phosphodiester groups of nucleic acids and are therefore a good indication of nucleic acid concentration in the samples. The presence of

these bands and lack of other spectral features is to be expected in basal cells due to the high nucleus to cytoplasm ratio.

In addition to the peaks discussed, epithelial cells gave spectra which showed 3 distinctive peaks at 1152, 1079, 1036cm^{-1}. These peaks result from glycogen and have been called the "glycogen triad" (Romeo *et al.* (2002a)). Chang *et al.* (2003) also recorded spectra from basal cells. However, the spectra presented in that study bear a remarkable resemblance to the spectra recorded from connective tissue in this present study particularly in terms of overall band shape and peaks at 1204 and 1283cm^{-1}. The connective tissue spectra seen in the present study correlate with the results presented by Chiriboga *et al.* (1998) and Wong *et al.* (1995).

Differences were seen in the spectra of epithelial cells as they matured and migrated towards the superficial layer. It was possible to see increasing glycogen concentrations by increases in the relative intensities of the glycogen triad. Again, Amide I bands were seen to shift indicating changes in protein structure.

The IR spectra from 10 points recorded from normal epithelial tissue from a single sample compare well to previously published spectra of normal cervical epithelium as seen by Wong *et al.* (1991), Mordechai *et al.* (2004), Romeo *et al.* (2002a), Mantsch *et al.* (2002), Wong *et al.* (1995), Diem *et al.* (2004), Diem *et al.* (2002), Wong *et al.* (2002), Chang *et al.* (2003), Cohenford *et al.* (1998), Fung *et al.* (1997), Neviliappan *et al.* (2002), Sindhuphak *et al.* (2003), Wood *et al.* (2003) and Chiriboga *et al.* (1998a). There have been no publications of spectra of normal cervical epithelium that dispute the spectra seen in the

present study. The only difference between results presented in the present study and those seen by other groups, is the constant shifting of the Amide I band. Some authors have not indicated Amide I bands in their results and it is therefore not possible to see whether there are Amide I band shifts present (Chang *et al.* 2003, Sindhuphak *et al.* 2002, Fung *et al.* 1997, Wong *et al.* 1995 and Morris *et al.* 1995). Of the studies where the Amide I band is presented (Wong *et al.* 2002, Wong *et al.* 1991, Mordechai *et al.* 2004, Mantsch *et al.* 2002, Diem *et al.* 2004, Cohenford *et al.* 1998, Neviliappan *et al.* 2001, Wood *et al.* 2003, Romeo *et al.* 2004, Romeo *et al.* 2002a), there appears to be less Amide I band shifting than is seen in the present study. It has been suggested that this shifting is due to the "dispersion artefact" as highlighted by Max Diem's group (Mohlenhoff *et al.* 2005). As samples were recorded in reflection mode, differences in sample thickness result in different scattering properties. This "non-Mie" type scattering results in the tissue spectrum being superimposed on an undulating background which results in a shift in the Amide I band (and to a lesser extent the Amide II band). Similar problems have been encountered by other groups, who wherever possible minimised the problem by recording spectra from the thinnest extremities of the sample (Boydsten White, personal communication).

Apart from shifts in both the Amide I and II bands, spectra of the normal epithelial tissue display consistent peaks at 1460, 1397, 1236, 1152, 1079 and 1036cm^{-1}. The glycogen triad (1152, 1079 and 1036cm^{-1}) are clearly present at each of the points sampled.

The interpatient variability was also examined and it was unsurprising to find shifting in the Amide I (and to a lesser extent Amide II) bands. Also a number of samples displayed residual wax contamination with bands at 1462 and 1472cm^{-1}. Interestingly, on close

examination of the spectra produced by other studies that examined dewaxed FFPP sections (Wood *et al.* 2004, Mordechai *et al.* 2004), small narrow peaks can be seen in this region also. The interpatient variability shows promising results. The asymmetric and symmetric stretching of phosphate (1236 and 1079cm^{-1}) from nucleic acids are present in all spectra. The glycogen triad is present in all spectra, although changes in levels of absorbance indicate different concentrations of glycogen in different samples. Apart from biological variation, other possible explanations for this include the point in the menstrual and ovarian cycles, as well as effects of oral contraceptives. Glycogen is accumulated in the intermediate epithelial cells as a result of estrogenic stimulation (Romeo *et al.* 2002b). The comparison of the spectra obtained from the different samples also highlighted a band at 930cm^{-1}. This band was seen to increase in intensity along with the glycogen triad. This band has been noted only once previously (Neviliappan *et al.* 2002) where it was assigned to nucleic acids. The present study would suggest that the band is in fact due to a pyranose asymmetric ring vibration of glycogen and not nucleic acid vibration as previously suggested. This suggestion is supported by the fact that pyranose rings are found in sugars (including glycogen) and the band was found to increase in intensity wherever the glycogen triad increased. Where the glycogen triad was absent, the band at 930 cm^{-1} was also absent.

The spectra obtained from invasive carcinoma areas of a single patient were examined, with a view to evaluating the interpatient variability within a single patient. It was immediately clear that the spectra displayed different characteristics to that of normal epithelial tissue. The Amide I band displayed only a small shift of ≈5cm^{-1} and was centred around 1660cm^{-1}. The most distinctive features of the spectra are the two broad bands at 1236 and 1079cm^{-1}

which have been previously assigned as phosphate stretching of nucleic acids. A new band was also seen at 970cm^{-1}, which has also been attributed to vibrations of nucleic acids and proteins (Wong *et al.* 1995). The amide III band could theoretically contribute to the intensity of the PO_2^- band at 1236cm^{-1}. However this can be ruled out for two reasons. The frequency of the Amide III band depends on the secondary structure of proteins, which also is determined by the position of the Amide I band. For example, Figure 5.10 shows the Amide I band at 1660cm^{-1} and thus the proteins are predominantly α- helical. The Amide III band of proteins with α- helical structures is located at 1260-1290cm^{-1} and therefore does not contribute to the intensity of the PO_2^- band at 1236cm^{-1} (Wong *et al.* 1991). Also in previous samples, were the 1236cm^{-1} band to be the result of Amide III and not PO_2^- vibrations, the band would shift as the Amide I band shifted and this was not observed. The spectra of invasive carcinoma shows a great deal of similarity to the spectra recorded from basal cells. This is to be expected due to the high proliferation rates associated with cancer cells.

It is also evident from Figure 5.11 that there are no glycogen contributions associated with spectra of invasive carcinoma. Glycogen depletion has been seen in all previous studies. Glucose is stored in the form of glycogen in cells for use as an oxidisable energy source. It has been suggested that glycogen depletion may be as a result of the increased energy required by proliferating cells.

A good degree of consistency can be seen among samples when the interpatient variability among carcinoma samples was examined, although Amide I band variation between

samples was evident. It is clear that the variation in the Amide I band does not result in shifts in the band at 1236cm^{-1}, further confirming the Amide III hypothesis presented earlier.

The spot to spot variation and interpatient variability of CIN tissue were examined. The spectra of the tissue bore a marked resemblance to the spectra of invasive carcinoma, though the intensity of the nucleic acid bands was somewhat reduced. Other studies investigating CIN tissue have had varying results. This is perhaps unsurprising given that CIN tissue by its nature is difficult to classify into a grade. Some groups showed glycogen contribution present in the spectra of neoplastic cells (Romeo *et al.* 2002a, Fung *et al.* 1997, Morris *et al.* 1995), while others did not (Chang *et al.* 2003, Diem *et al.* 2002, Wood *et al.* 2003, Mordechai *et al.* 2004, and Neviliappan *et al.* 2002). In the present study there was no evidence of glycogen, although the spectra were recorded from the epithelial layer.

When directly compared to the spectra of normal epithelial tissue, the spectra recorded from invasive carcinoma showed a marked difference. Results presented compared well to previously published reports, particularly those pertaining to dewaxed FFPP cervical tissue sections. The present study found a dramatic reduction in glycogen band intensity (1152, 1079 and 1036cm^{-1}), as well as an increase in the nucleic acid band intensity (1236cm^{-1}) with the onset of carcinogenesis. Unfortunately no consistent changes could be seen in either the Amide I or II bands due to the dispersion artefact discussed previously.

Synchrotron Fourier transform infrared spectroscopy has been shown to be an excellent method of exploring the biochemical composition of different cell types found in human

tissue. It has been shown that spectral characteristics of these different cell types can be used to differentiate between them. FTIR spectroscopy has also been shown to differentiate between normal, cancerous and pre-cancerous tissues. It has highlighted some of the biochemical markers associated with cancerous tissue. The dramatic reduction in glycogen levels and the increase in nucleic acids with the onset of carcinogenesis has been clearly demonstrated. It has been shown that dewaxed FFPP sections can indeed be utilised in a diagnostic capacity, to distinguish between invasive carcinoma and normal epithelial cells and also to highlight pre-malignancy. Peaks identified as diagnostically significant, have not included any of the peaks previously outlined as being affected during the processing and dewaxing stages. FTIR measurements have also been carried out on a conventional benchtop spectrometer (although with larger spatial resolution) and spectra compare favourably with results presented in the present study (results not shown). The spatial resolution associated with SR-FTIR spectroscopy is an advantage for direct correlation with Raman spectroscopic measurements, which are presented in the following Chapter.

Chapter 6

Raman spectroscopy of cervical tissue

6.1 Introduction

Chapter 5 showed how FTIR spectroscopy could be used to distinguish between the biochemical signatures of normal and abnormal cervical tissue. The most apparent differences observed included very strong glycogen contributions in normal tissue, while invasive carcinoma tissue showed no glycogen and increased nucleic acid concentrations. The current Chapter aims to use Raman spectroscopy on the same samples investigated in Chapter 5. It is hoped that the increased spatial and spectral resolution afforded by Raman spectroscopy could yield additional spectral markers unseen because of the broad bands associated with FTIR spectroscopy.

Many groups have investigated the use of Raman spectroscopy for cancer diagnosis (Section 1.5). Each of these studies has employed its own experimental procedure and instrument settings. In the study of human skin (for example, Gniadecka *et al.* 1997, Knudsen *et al.* 2002), near-infrared (NIR) laser light (1064nm) was used as an excitation wavelength, whereas, Caspers *et al.* (2001), Nijssen *et al.* (2002) and Schutt *et al.* (2002), also examining human skin, employed both 730nm and 850nm excitation. This was due to optimisation of the quantum efficiency of the charged coupled device (CCD) detectors being used. In other studies on brain tissue, Lakshmi *et al.* (2002), Koljenović *et al.* (2002) and Mizuno *et al.* (1994), used excitation wavelengths of 785, 847 and 1064 nm respectively. All of the Raman studies also handled tissues in different ways (dewaxed histological sections, cryosections, tissue pieces refrigerated in PBS, sample collected by

biopsy/ from autopsy, recorded in solution, dried) and used many different recording parameters (different objective lenses, accumulation times, resolution etc.).

In relation to the studies on cervical tissue using Raman spectroscopy, Mahadevan-Jansen *et al.* (1998) and Utzinger *et al.* (2001) used 789nm excitation for both *in vitro* and *in vivo* studies. Liu *et al.* (1992) used 1064nm excitation on the cervical tissue examined, while Yazdi *et al.* (1998) used 257 nm laser excitation to examine cultured cervical cells.

Throughout the present study 514.5 nm laser excitation has been used. As seen in Chapters 3 and 4, both cervical and parenchymal tissue were examined and both yielded good quality Raman spectra. Any background has been easily removed by using a simple baseline subtraction.

Although Raman spectroscopy has many advantages in relation to possible *in-vivo* detection of cancer, there has only been a limited amount of research into cervical samples.

Liu *et al.* (1992) conducted an early study into Raman, fluorescence and time-resolved light scattering as optical diagnostic techniques on a selection of gynaecological tissues. A small portion of the study dealt with the Raman spectral properties of normal and malignant cervical tissue. Samples were examined using a 1064 nm laser and no sample preparation was carried out. In malignant tissue two Raman peaks were seen at 1657 and 1445cm^{-1} where the intensity $I_{1657} < I_{1445}$. Benign and normal tissues were found to have common peaks at 1659, 1445, 1262 and 934 cm^{-1}. The study found the relative intensity of the 1659 cm^{-1} band to the band at 1445cm^{-1} in normal tissue was reversed (i.e. $I_{1659} > I_{1445}$). A broad

band at 1262 cm^{-1} was observed in normal and benign tissue, but was absent in the spectra of malignant tissue. The study also highlights a peak at 934 cm^{-1} present in spectra of normal, while absent in spectra of malignant tissue.

Mahadevan-Jansen *et. al* (1998) conducted a study on 18 patients investigating the potential of precancer diagnosis. Near infrared (789nm) Raman spectra were recorded from 36 snap frozen biopsies. Raman peaks were present in spectra of all samples and were observed at 626, 818, 978, 1070, 1246, 1330, 1454 and 1656 cm^{-1}. Fischer discriminate analysis was used to differentiate between normal, metaplasia, inflammation, low-grade squamous intraepithelial lesion (SIL) and high-grade SIL. In general peaks attributed to collagen (1656, 1070 cm^{-1}) decrease in intensity in spectra of tissues with SIL, while peaks attributed to phospholipids, DNA and glucose 1-phosphate (1454, 1330, 978 cm^{-1}) increase in intensity in spectra of tissues with SIL. This was deemed consistent with the biology of the neoplastic process in which the number of cells in the epithelium increases as SIL develops. Using multivariate methods precancers could be differentiated from all other tissue with a sensitivity of 82% and specificity of 92% in an unbiased test.

Yadzi *et. al* (1998) investigated resonance Raman spectroscopy (257nm) of normal and malignant cultured breast and cervical cells in PBS suspension. Cell spectra obtained closely resembled that of DNA with peaks around 1330, 1480 and 1580 cm^{-1}, due to the nucleotide bases. The intensity ratios of Raman spectral peaks 1480/1614 cm^{-1} and 1480/1540 cm^{-1}, which are sensitive to the concentration of nucleic acids relative to cell proteins, were found to be higher in malignant cells than in normal cells. Normal and malignant cells could also be differentiated by using the ratio at 1330/1480 cm^{-1}, which

may be the result of RNA concentration (however this observation was not as pronounced in cervical cells).

Utzinger *et. al* (2001) evaluated the potential of near-infrared (789 nm) Raman spectroscopy for *in-vivo* detection of cervical precancers. Twenty-four measurements were made in 13 patients. Spectra were recorded from one normal and one abnormal area of the cervix. Raman spectra contained peaks at 1070, 1180, 1195, 1210, 1245, 1330, 1400, 1454, 1505, 1555, 1656 and 1760 cm^{-1}. As the diagnosis progressed from normal to inflammation to metaplasia and through squamous dysplasia, the intensity of the Raman peak at 1454 cm^{-1} increased. The intensity of the Raman peaks at 1330 and 1656cm^{-1} was greater for abnormal samples than normal samples. The ratios of intensities at 1454/1656 cm^{-1} was greater on average for squamous dysplasia samples than all other tissue types, while the ratio of intensities at 1330/1454 cm^{-1} was lower on average for samples with squamous dysplasia than all other tissue types. Spectra measured *in-vivo* were found to resemble those measured *in-vitro*.

Before examining the spectroscopy of cancer tissue, the spectral properties of the different cell types of cervical tissue were examined. As outlined in Chapter 1, the main cell types in cervical tissue are basal cells, epithelial cells (at different stages of maturity) and connective tissue. The properties of epithelial cells are of specific interest as it is epithelial cell proliferation that takes place in cancerous growths. The spectral properties of these cells as they mature, as well as the surrounding cell types, should be understood for a diagnostic algorithm to be developed. Spectra were recorded from different cell types from different patients and the spectra were compared to evaluate both spot to spot variation

within a single patient, as well as interpatient variability. Once this variation had been examined, spectra were recorded from normal cervical tissue and compared to the spectra obtained from invasive carcinoma, as well as neoplastic tissue. This chapter seeks to qualitatively examine spectral differences between classes of cells and normal and abnormal cervical tissue. As seen in Chapter 5, SR-FTIR spectroscopy proved capable of differentiating these cell / tissue types. These results will assist the present study, before an automated process is applied to the datasets.

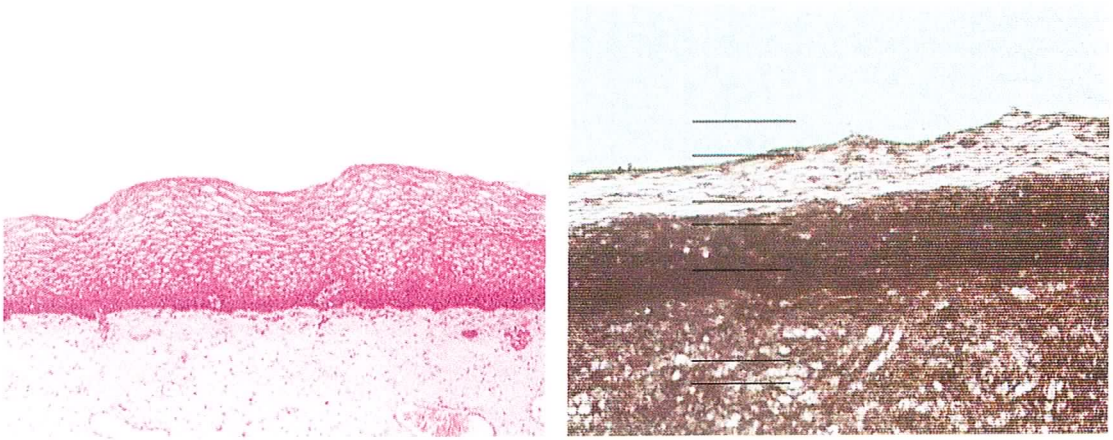
6.2 Materials and Methods

6.2.1 Instrumentation

The Instruments S.A. Labram 1B Raman spectroscopic confocal microscope was used, with an Argon Ion laser operating at a wavelength of 514.5nm, as described in section 3.2.1. Spectra were recorded for 150 seconds and measured over a spectral range of 400 to 1900 cm^{-1}. Once acquired each spectrum was baseline corrected, dark current subtracted and lightly filtered using a third order linear model to improve clarity.

6.2.2 Sample preparation (FFPP sections)

20 normal and 20 invasive carcinoma (including 10 CIN) were prepared as outlined in section 5.2.2. Unstained tissue sections were examined under the Raman microscope (Figures 6.1 and 6.2).



(A) (B)

Figure 6.1 (A) Haematoxylin and Eosin stained tissue section and (B) unstained tissue section, x50 objective

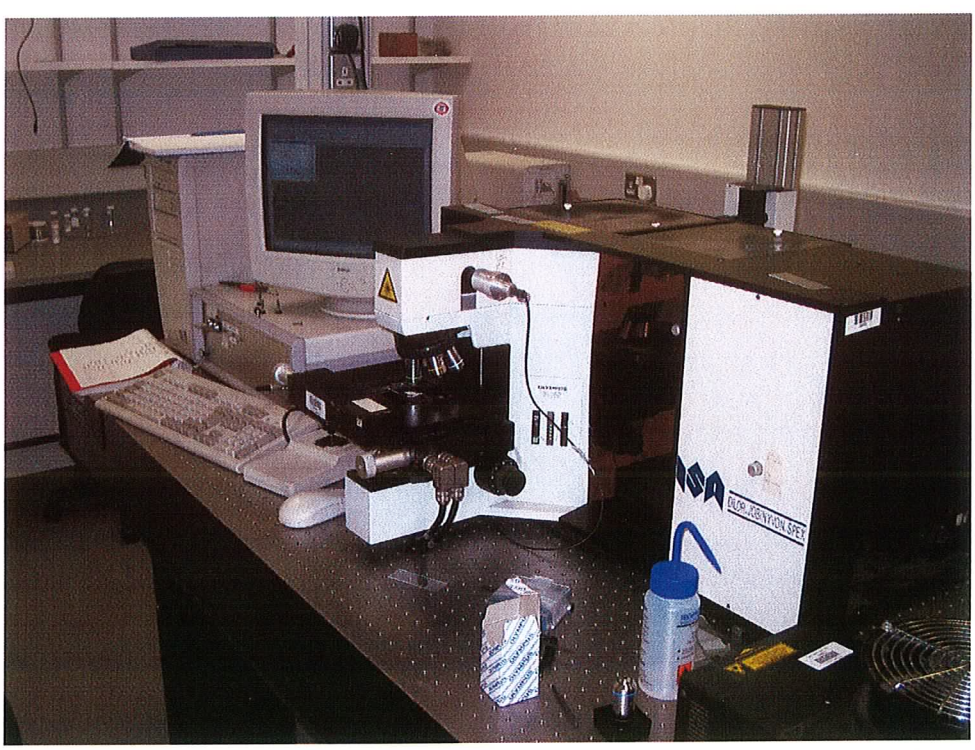


Figure 6.2 Raman spectrometer and microscope with tissue sample on stage

6.3 Results

6.3.1 FFPP Normal Cervical Tissue

The Raman spectra of basal cells, epithelial cells and connective tissue were recorded from a number of different samples. Figure 6.3 shows the spectra from each cell type over a large spectral range (300 – 4000 cm^{-1}). It is clear that there are subtle differences between the three different cell types and that there are two main regions of Raman activity associated with the samples. The first region (400 – 1900 cm^{-1}), also known as the fingerprint region, contains the majority of the structural information relating to the sample i.e. amide bands, carbon-carbon, CH_2 and ring vibrations. The vibrations in the 3000 cm^{-1} region are mainly due to CH_2 and CH_3 stretching vibrations. This region was found to vary greatly from sample to sample and as seen in Figure 6.2 the region saturated well before the spectral features in the fingerprint region had reached their maximum intensity. This region is also associated with lipid activity and lipid structure is known to be affected by formalin fixation (Shim *et al.* (1996)). Also, to achieve the greatest spectral resolution, it was desirable to record spectra using the 1800 grooves/mm grating. To cover the entire spectral range using this grating would have meant each spectrum would take 3 times as long to record, slowing down the recording process and increasing the likelihood of burning the sample under investigation. For all of the above reasons it was decided to focus attention on the fingerprint region alone.

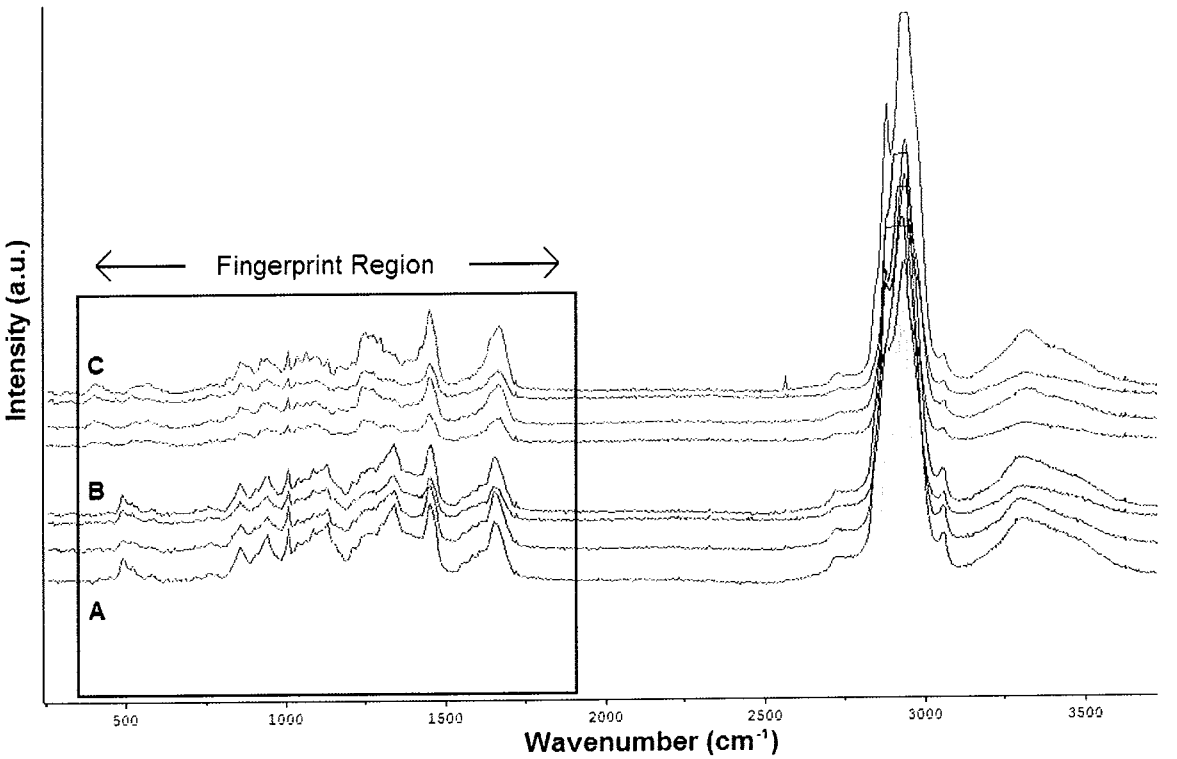


Figure 6.3 Raman spectra of a second tissue sample from A) basal cells B) epithelial cells and C) connective tissue. The "fingerprint region" (400 – 1900 cm^{-1}) has also been highlighted.

Figure 6.4 shows the fingerprint region in more detail. Spectra were recorded from a single sample and each spectrum represents a different spot within the sample. The spectra of the three different cell types do have a degree of similarity (ring breathing, CC stretch, Amide III, CH bending and Amide I bands). This has been seen previously in Chapter 2 when different tissue types were examined (Figure 2.16). However there are also many spectral features differentiating the different cell types. The spectra of basal cells (Fig, 6.4A) show strong bands at 724, 779 and 1578 cm^{-1} which are characteristic of nucleic acids. Spectra of epithelial cells (Figure 6.4B) have characteristic glycogen bands at 482, 849, 938, 1082 and

1336 cm^{-1}. Collagen contributions are clearly seen in the spectra of connective tissue (Figure 6.4C) at 850, 940 and 1245 cm^{-1}.

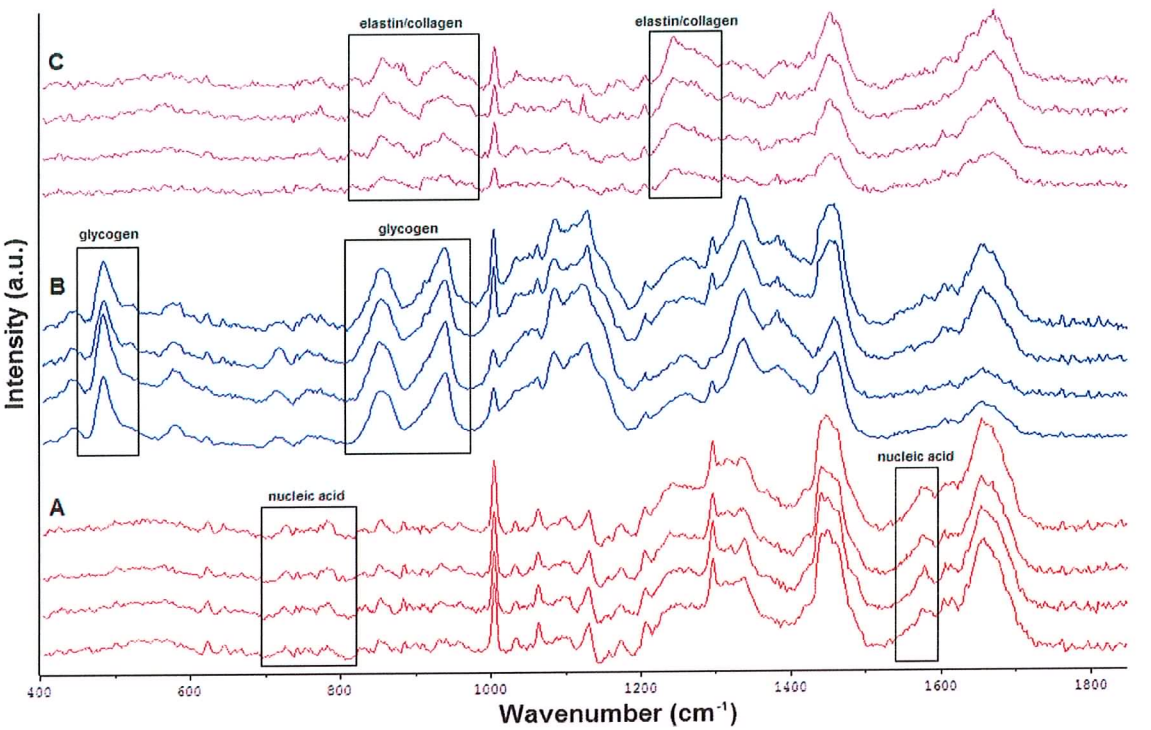


Figure 6.4 Raman spectra recorded from fingerprint region with 1800 grooves/mm grating

A) basal cells B) epithelial cells C) connective tissue

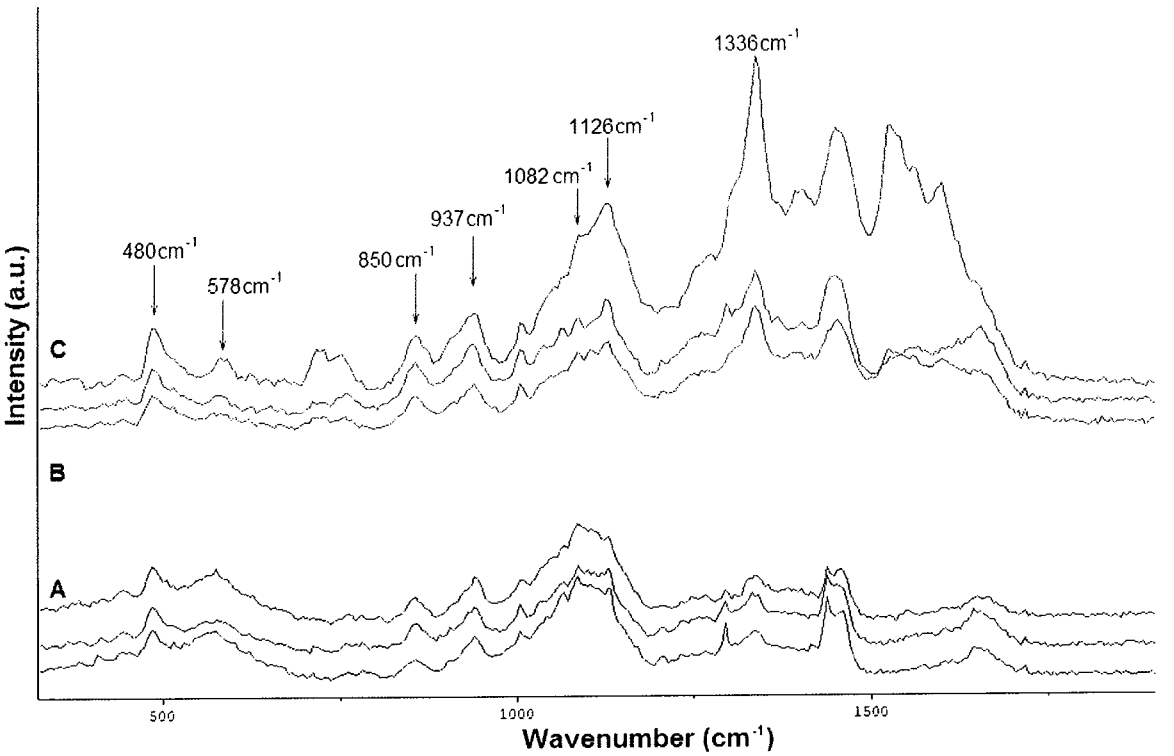


Figure 6.5 Raman spectra of A) parabasal, B) intermediate and C) superficial epithelial cells, recorded with 1800 lines/mm grating

Figure 6.5 examines the spectral properties of the different types of epithelial cells, namely, parabasal, intermediate and superficial squamous epithelial cells. It was found that the glycogen levels increased as the cell matured. This accumulation is visible by an increase in the relative intensity of the spectral assignments associated with glycogen, namely those bands at 480, 578, 850, 937, 1082, 1126 and 1336 cm^{-1}. In order to maintain consistency in the type of epithelial cell being sampled and with a view to diagnosing cancer at an early stage (neoplasia), a decision had to be taken as to which category of squamous epithelial cell was to be sampled. Should neoplastic growth commence, it is the intermediate epithelial cells that will take on the abnormal characteristics associated with neoplasia

(CINII). Coupled with the fact that the overwhelming majority of cells collected during a cervical smear are superficial or intermediate epithelial cells, it made intermediate cells the natural choice. This leaves open the possibility that understanding gained from studying the spectral properties of tissue sections could be applied to cervical smear cells at a later date.

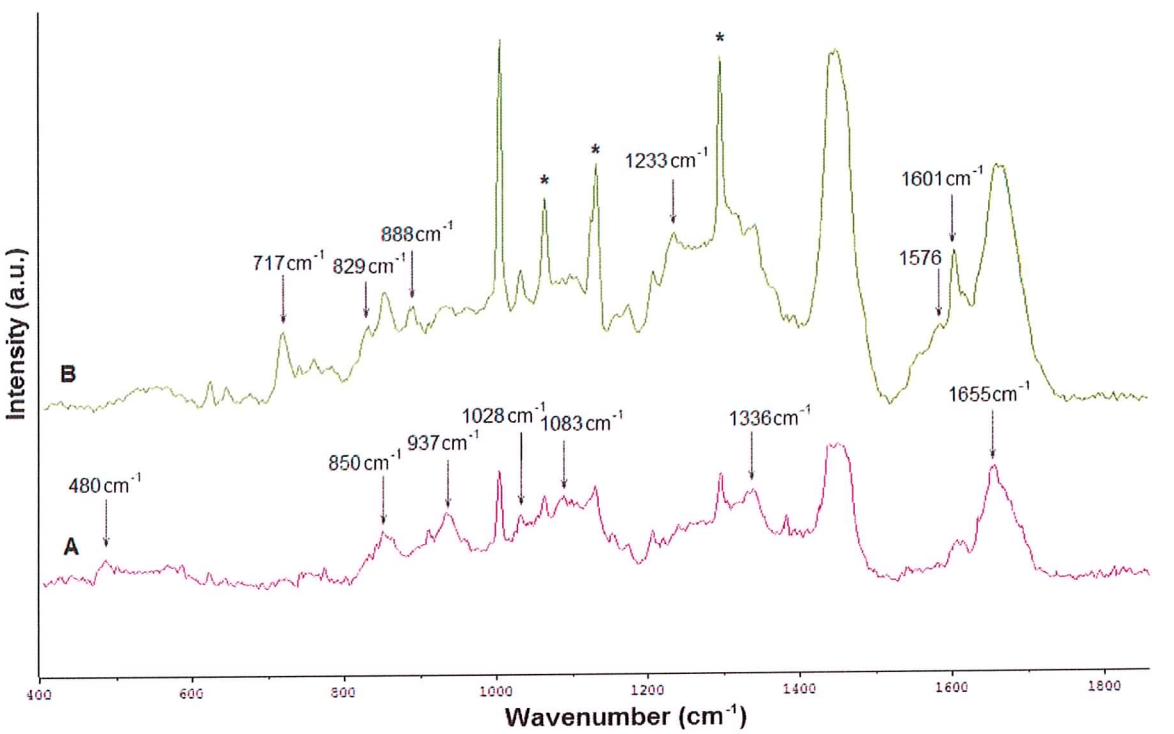


Figure 6.6 Raman spectra of A) normal epithelial and B) columnar cells

As cervical columnar cells from the endocervical canal may also be gathered during a cervical smear, a comparison was made between normal epithelial cells and columnar cells (Figure 6.6). Columnar cells exhibited different spectral properties from epithelial cells, as expected from previous FTIR studies (Wong *et al.* 2002). The most noticeable differences included an absence of glycogen from columnar cells (evident from the absence of peaks at

480, 850 and 937 cm^{-1}), as well as strong contributions at 717 cm^{-1} and 1601 cm^{-1} (CH_2 rocking and CO stretch; C=C stretch respectively). Spectra of columnar cells also exhibited peaks at 829, 888, 1233 and 1576 cm^{-1}, corresponding to O-P-O stretch, C-C stretch, Amide III, Adenine (A) / Guanine (G) modes respectively (Notingher *et al.* 2003). Spectra of columnar cells also exhibited much stronger wax contributions than seen in spectra of epithelial cells (peaks highlighted by asterisks in Figure 6.6).

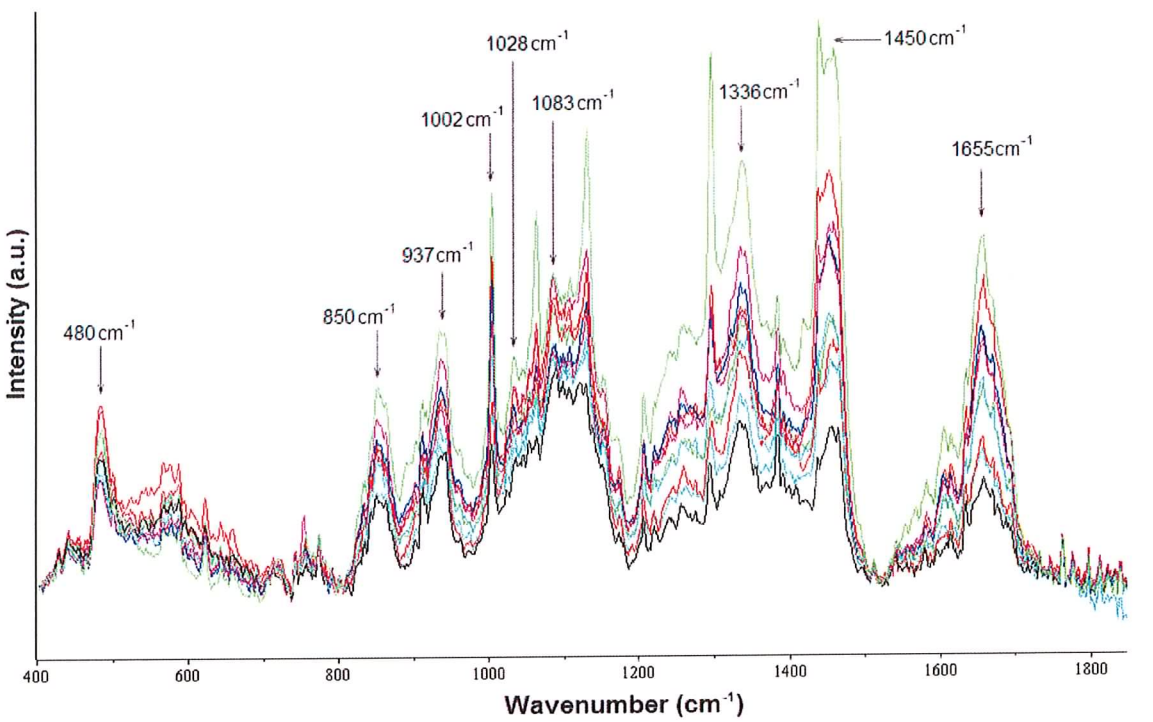


Figure 6.7 Raman spectra of 10 different intermediate squamous epithelial cells from a single patient

Figure 6.7 shows 10 Raman spectra collected from the intermediate squamous epithelial region of a single sample. The 10 spectra have all of the same bands, giving a good representation of the tissue. Although there is some variation in the overall intensity of the

spectra, the relative intensities remain constant. Some of the features of interest have been highlighted and include;

480 cm^{-1}	skeletal ring vibration
850 cm^{-1}	δ (CCH) aromatic
937 cm^{-1}	υ (CC)
1002 cm^{-1}	υ (CC) aromatic ring
1028 cm^{-1}	υ (CC) skeletal conformation
1083 cm^{-1}	υ (CC) skeletal random
1336 cm^{-1}	δ (CH_2)
1450 cm^{-1}	δ (CH_2) proteins and lipids
1655 cm^{-1}	υ (C=O) amide I (α – helix)

As seen in the spectra shown in Figure 6.5, glycogen features very strongly in the spectra of intermediate epithelial cells seen in Figure 6.7 as would be expected. Figure 6.8 displays the same spectra as Figure 6.7, however the spectra have been offset to improve clarity.

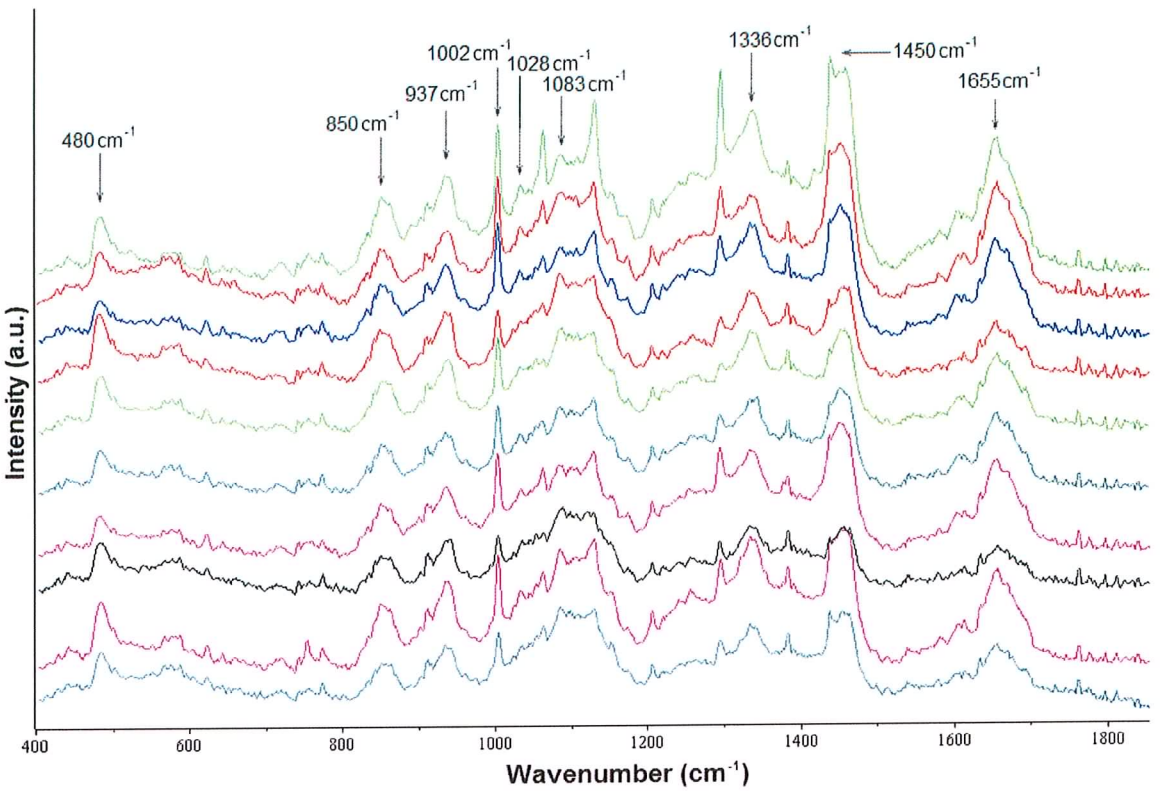


Figure 6.8 Raman spectra of 10 different epithelial cells from a single patient offset to improve clarity

Figure 6.9 examines a single Raman spectrum taken from 10 different samples, in order to evaluate the interpatient variability. Again, all spectra were recorded from the intermediate epithelial cells. The main features are common to all of the different samples examined. The variability seen among different patients will be examined further in Chapter 7.

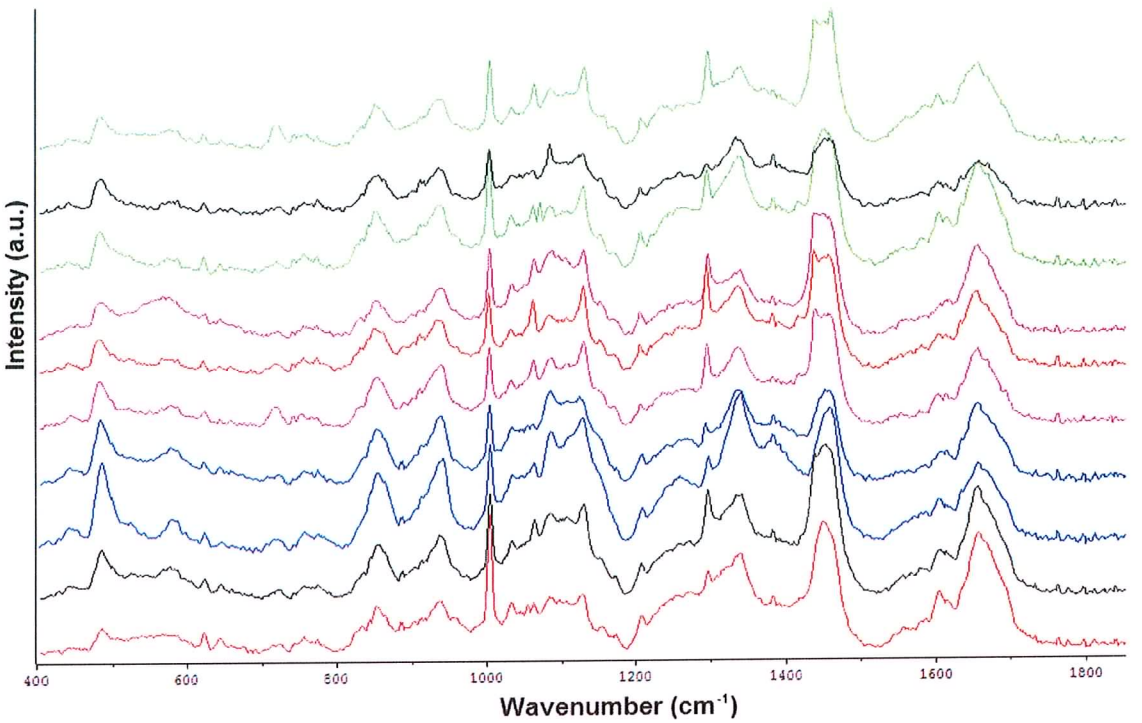


Figure 6.9 Raman spectra of epithelial cells from 10 different patients to examine interpatient variability

6.3.2 FFPP Invasive Carcinoma

Areas of invasive carcinoma were marked on the Haematoxylin and Eosin (H+E) stained reference slide by a pathologist. Raman spectra were subsequently recorded from these areas on the unstained slide. Figure 6.10 shows 10 Raman spectra from a single sample.

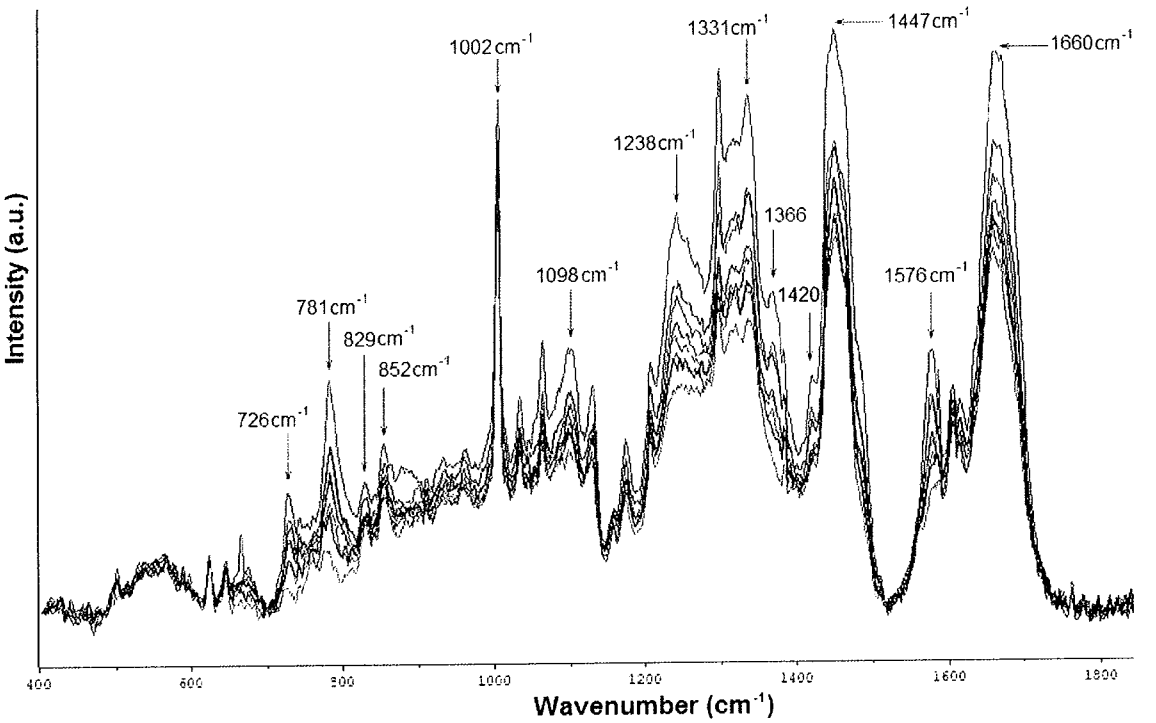


Figure 6.10 Raman spectra recorded from an invasive carcinoma region of a single patient

Again, the 10 spectra have all of the same bands, giving a good representation of the tissue. Although there is some variation in the overall intensity of the spectra, the relative intensities remain relatively constant. Some of the features of interest have been highlighted and include;

726 cm^{-1} nucleic acid base Adenine (A)

781 cm^{-1} nucleic acid bases, Cytosine (C), Tyrosine (T) ring breathing

829 cm^{-1} O-P-O DNA backbone

852 cm^{-1} δ (CCH) aromatic

1002 cm^{-1} υ (CC) aromatic ring

1098 cm^{-1} PO_2^-, DNA backbone

1238 cm^{-1} CH_2 deformation; Amide III

1331 cm^{-1} CH_2CH_3 wagging of polynucleotide chain

1366 cm^{-1} CH_3 symmetric stretch

1420 cm^{-1} CH_3 deformation

1447 cm^{-1} CH_2 deformation

1576 cm^{-1} nucleic acid bases Guanine (G), Adenine (A)

1660 cm^{-1} υ (C=O) amide I (α – helix)

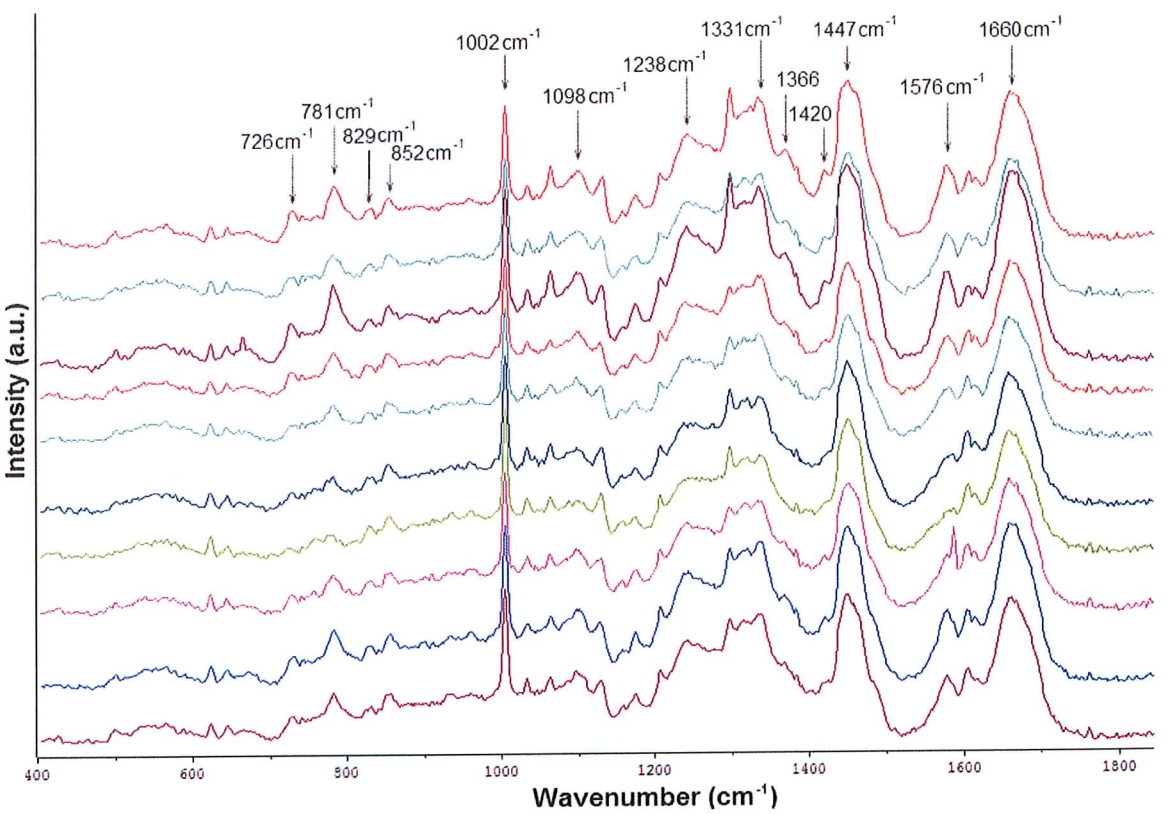


Figure 6.11 Raman spectra from an invasive carcinoma region of a single patient, offset to improve clarity

Figure 6.11 shows the same spectra offset. As seen with the normal tissue spot to spot variation does not pose a problem. Figure 6.12 examines the interpatient variability of 10 different samples. Again a high degree of repeatability was seen, although changes in the relative intensity of the peak at 1028cm^{-1} were observed.

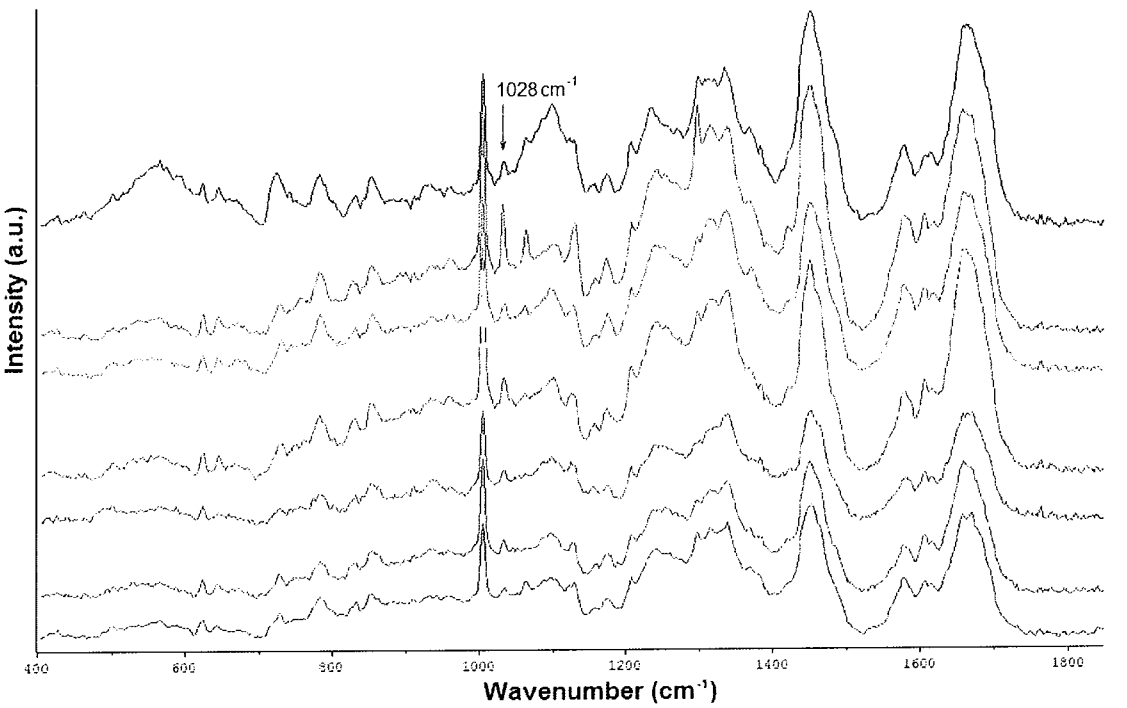


Figure 6.12 Raman spectra of invasive carcinoma recorded from 10 different patients to examine interpatient variability

Figure 6.13 compares the Raman spectra collected from normal epithelial cells and invasive carcinoma from a selection of different patients. Glycogen contributions are clearly visible in the spectra from the normal epithelial tissue (Figure 6.13A). The most obvious bands arise at 482 cm^{-1}, 849 cm^{-1} and 938 cm^{-1} and are due to glycogen skeletal deformation, CCH aromatic deformation and CCH deformation respectively. However there are also

other glycogen contributions not as apparent, including a CC stretching band at 1082 and CH_3CH_2 wagging at 1336 cm^{-1}.

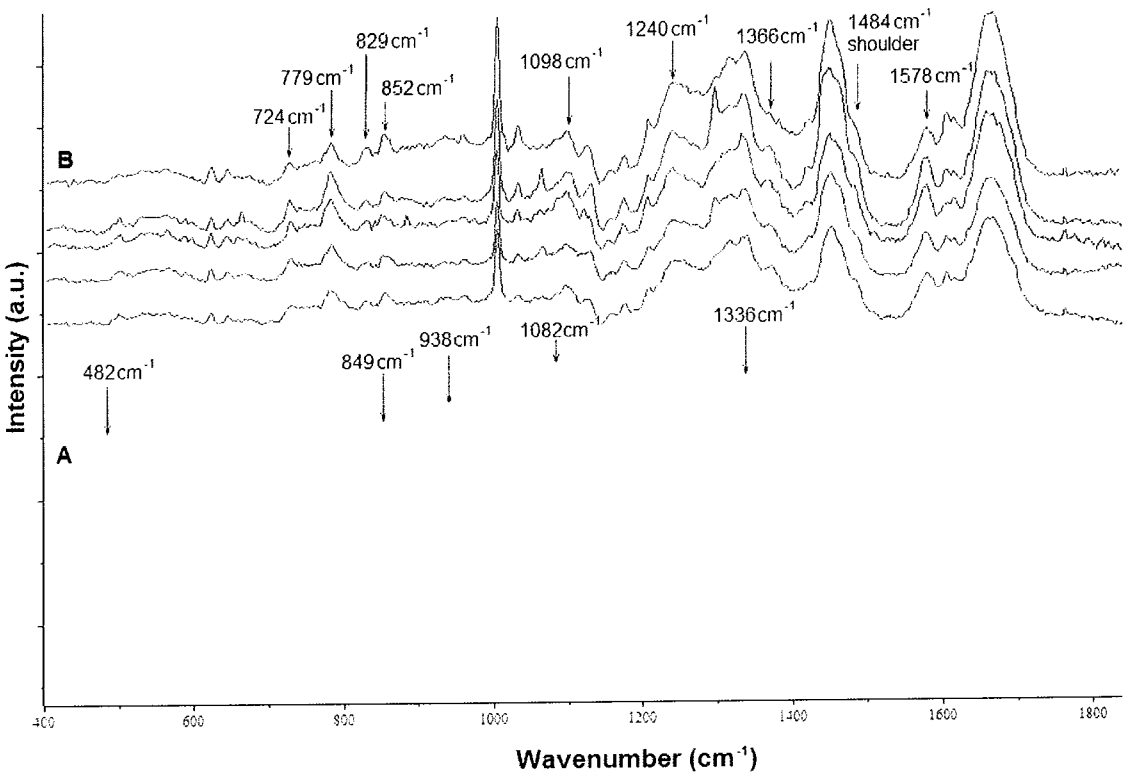


Figure 6.13 Raman spectra of A) normal epithelial tissue B) invasive carcinoma recorded from different patients

There is a dramatic reduction in intensity of these glycogen bands (482, 849 and 938 cm^{-1}) in the invasive carcinoma spectra, as well as a reduction in intensity of the CC stretching mode (1082 cm^{-1}).

The spectra of invasive carcinoma (Figure 6.13B) show many features of increased nucleic acid levels. These include prominent increases in intensity of bands at 724, 779 and 1578 cm^{-1}, but also increases in bands at 829, 852, 1002, 1098, 1240 cm^{-1}. Distinct bands were also seen at 1366 cm^{-1}, a shoulder at 1484 cm^{-1} and a band at 1578 cm^{-1}. An increase in the intensity of the Amide I band (1655cm^{-1}) was also observed in the spectra of carcinoma samples.

Spectra were recorded from glucose, glycogen, DNA and RNA. Figure 6.14 compares the spectra recorded from glucose (Figure 6.14(A)) and glycogen (Figure 6.14(B)), to that of normal epithelial tissue (Figure 6.14(C)).

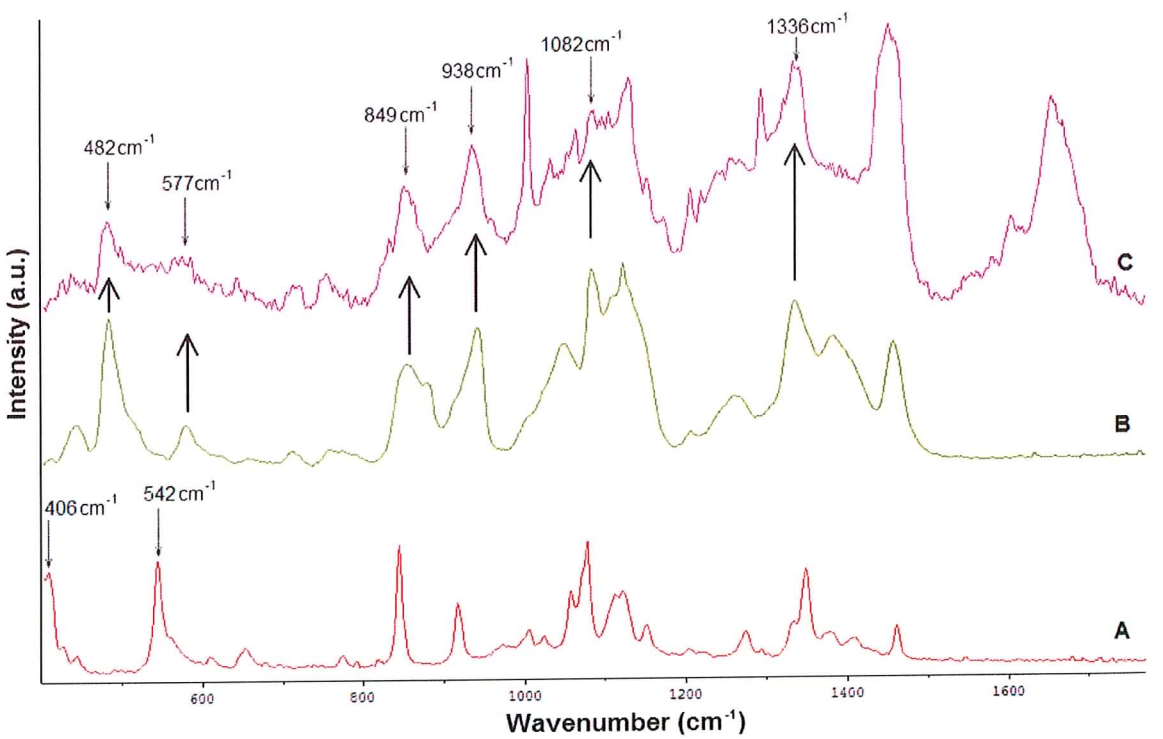


Figure 6.14 Raman spectra of A) glucose B) glycogen and C) normal epithelial cells

It is clear that the carbohydrate contributions seen in spectra of normal epithelial cells are due to glycogen and not glucose. Glucose exhibited two distinctive peaks at 406 and 542 cm^{-1} not present in spectra of glycogen. The spectrum of glycogen is very similar to that of normal cervical epithelium, with dominant peaks seen at 482, 577, 849, 938, 1082 and 1336 cm^{-1}.

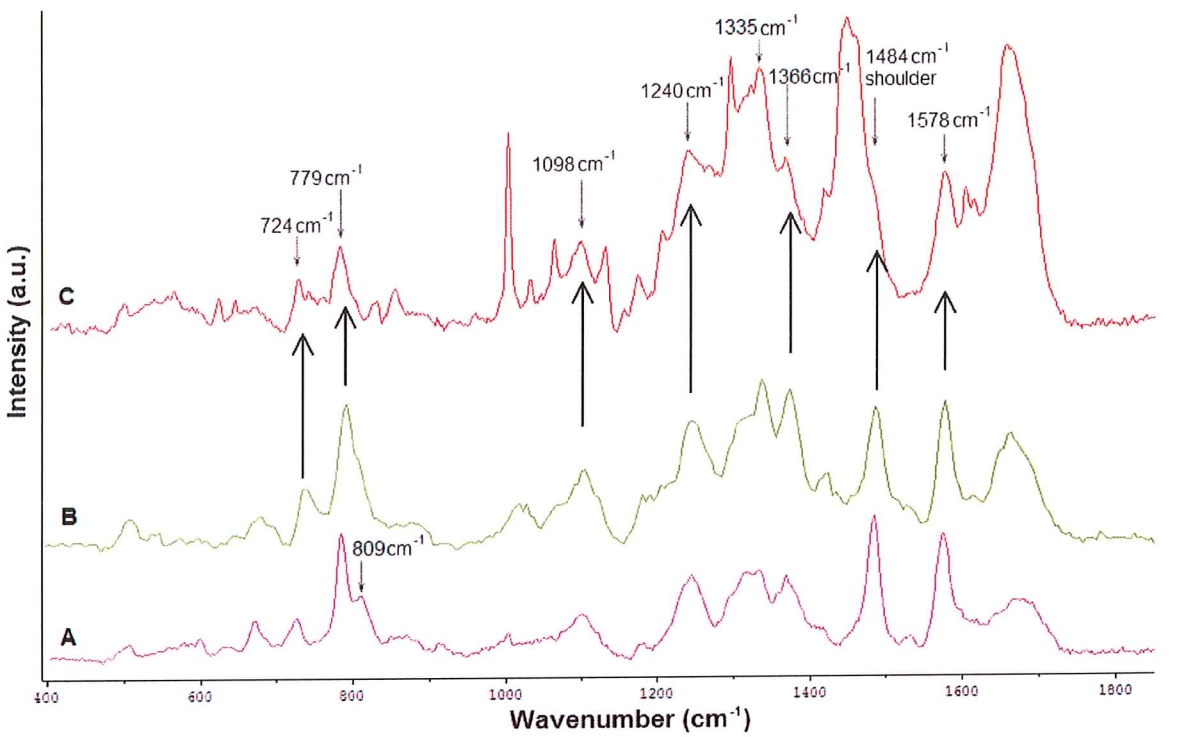


Figure 6.15 Raman spectra of A) RNA B) DNA and C) invasive carcinoma

Figure 6.15 compares the Raman spectra of DNA and RNA to the spectrum of invasive carcinoma. The comparison confirms the increase in DNA in the proliferating carcinoma cells. DNA and RNA can be distinguished by the presence of a double peak in spectra of RNA at 779 and 809 cm^{-1}, whereas spectra of DNA exhibit only a single peak at 779 cm^{-1}.

The presence of the Amide I band in the DNA spectrum (absent in that of glycogen), provides a possible explanation for the increase in Amide I vibrations seen in Figure 6.13.

6.3.3 FFPP Neoplastic Tissue

In order to examine whether or not pre-malignant changes could be highlighted using Raman spectroscopy, 10 areas of neoplasia (CIN III) were marked by a pathologist and the Raman spectra recorded from a single patient are shown in Figure 6.16.

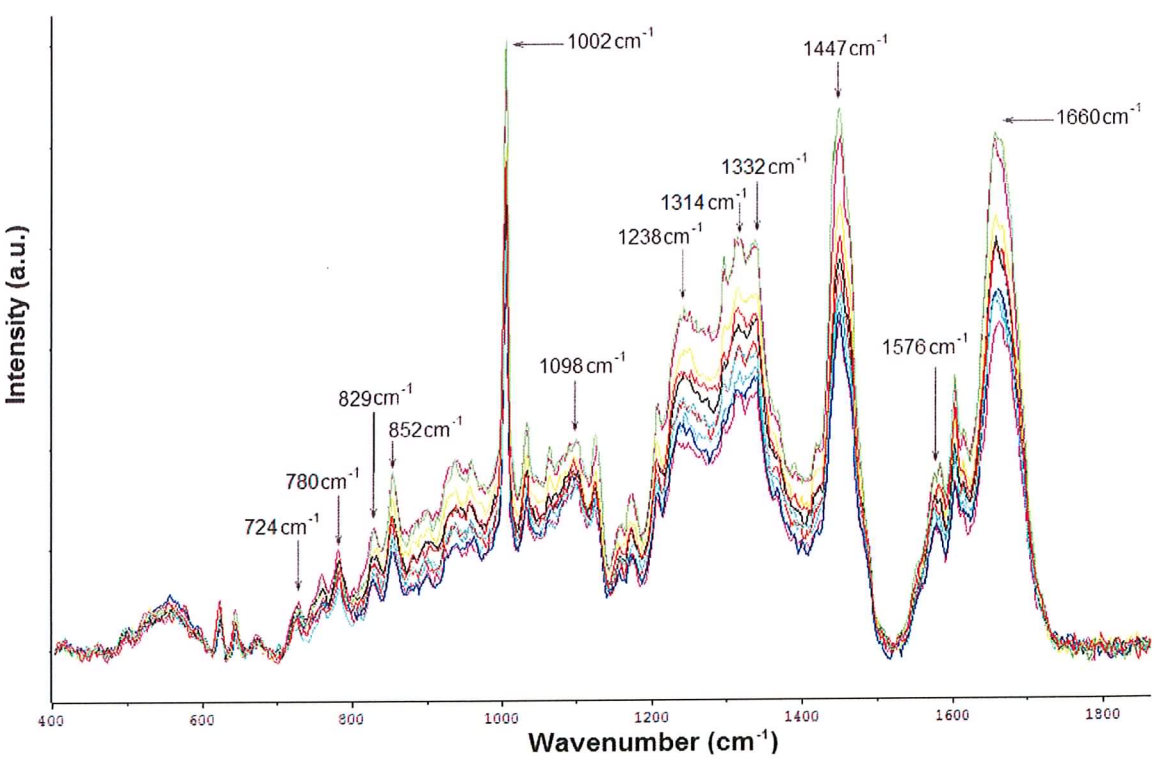


Figure 6.16 Raman spectra recorded from CIN region of a single patient

As seen with other tissue types, a good degree of consistency was seen from spot to spot. The spectra obtained bear a large degree of similarity to the invasive carcinoma spectra seen in Figure 6.10.

Peaks identified included;

724 cm^{-1}	nucleic acid base Adenine
780 cm^{-1}	nucleic acid bases, Cytosine and Tyrosine ring breathing
829 cm^{-1}	O-P-O DNA backbone
852 cm^{-1}	δ (CCH) aromatic
1002 cm^{-1}	υ (CC) aromatic ring
1098 cm^{-1}	PO_2^-, DNA backbone
1238 cm^{-1}	CH_2 deformation; Amide III
1314 cm^{-1}	CH_2CH_3 twisting
1332 cm^{-1}	CH_2CH_3 wagging of polynucleotide chain
1447 cm^{-1}	CH_2 deformation
1576 cm^{-1}	nucleic acid bases Guanine and Adenine
1660 cm^{-1}	υ (C=O) amide I (α – helix)

Figure 6.17 examines the same spectra from a single patient which have been offset to facilitate clarity. Figure 6.18 shows the spectra from 10 different samples and again a relatively good degree of consistency can be seen between samples.

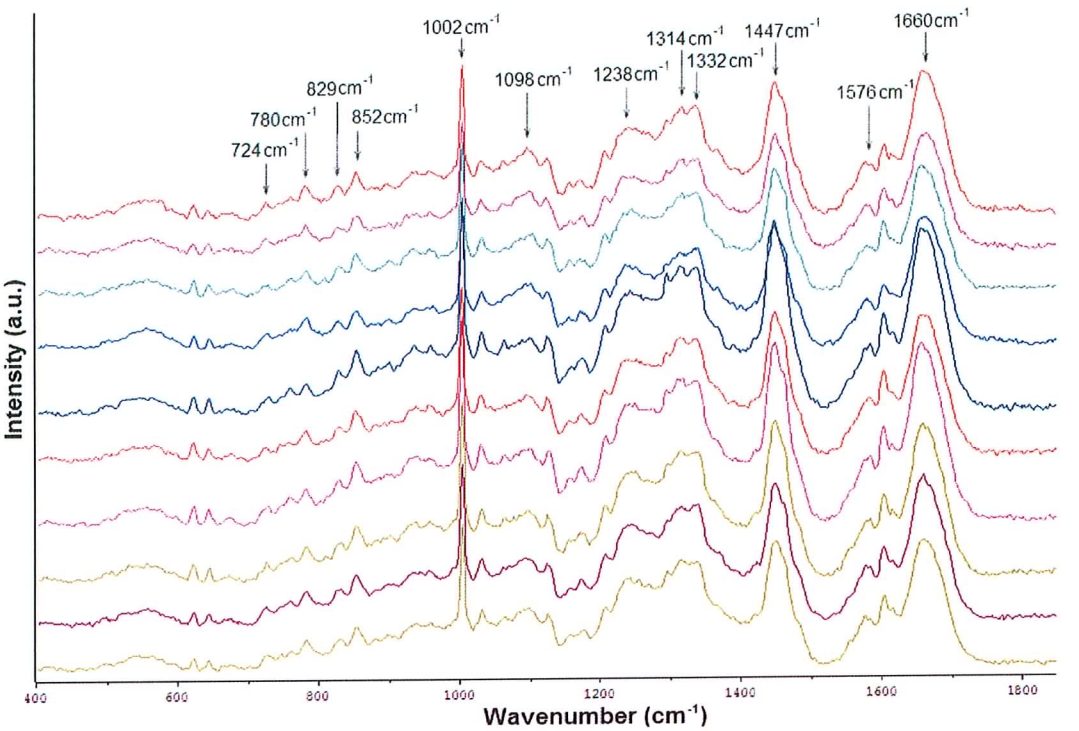


Figure 6.17 Raman spectra from CIN region of a single patient, offset to improve clarity

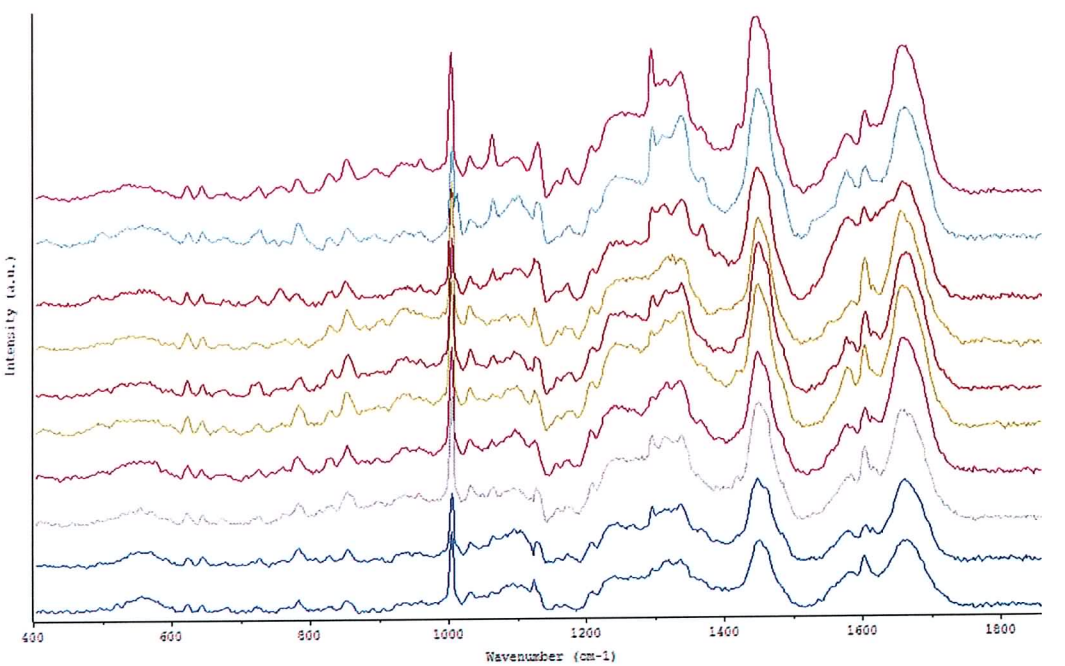


Figure 6.18 Raman spectra of CIN regions recorded from 10 different patients to examine interpatient variability

6.4 Conclusion

The results seen in this study have been very encouraging. Having selected a visible wavelength laser, the 1800 grooves/mm grating and choosing to examine the fingerprint region alone, the spectra obtained are of very good quality and contain an enormous amount of structural information about the sample under investigation. The small amounts of background encountered were easily removed by simple baseline subtraction. Problems encountered by other groups in relation to large fluorescence were found not to be a problem in the case of the FFPP cervical tissue sections examined. This may be due to fact that cervical tissue has relatively few endogenous fluorophores, unlike melanin in skin for example, which would add greatly to the levels of fluorescence.

Raman peaks were observed at 482, 578, 724, 779, 849, 938, 1002, 1030, 1060,1082, 1098, 1129, 1149, 1294, 1336, 1447, 1484, 1576, 1610 and 1660 cm^{-1} ($\pm 10 cm^{-1}$) in the spectra of normal cervical epithelium. The peaks can be used to provide information about both the presence and the concentration of various functional groups in order to develop an empirically based diagnostic algorithm to recognise cancer. This chapter aims to determine the diagnostic potential of each of these peaks in an empirical way.

Initially, the characteristics of different cell types seen in cervical tissue were examined (Figure 6.4). The spectra of basal cells showed strong bands at 724 cm^{-1}, 779 cm^{-1} and 1578 cm^{-1}. These bands are associated with nucleic acid vibrations, as seen by Kneipp *et al.* (2003a) in his study of breast duct epithelia. The same contributions were observed in the

spectra of salmon DNA (Figure 2.10). This high concentration of nucleic acid could be explained as follows. Firstly, the morphology of basal cells consists of a single line of tightly packed cells, with large nuclei in relation to the compacted surrounding cytoplasm. Secondly, these cells are constantly dividing, providing cells to the parabasal layer. For both of these reasons a high concentration of DNA is to be expected in basal cells. Epithelial cells showed high glycogen concentrations with spectra showing bands at 480, 578, 850, 937, 1082, 1126 and 1336 cm^{-1}. Similar results have been reported by Choo-Smith *et al.* (2001). Connective tissue is characterised by the abundance of intracellular material produced by its cells. This material is composed mostly of a network of tough protein fibres, collagen and elastin. Collagen (elastin has a very similar spectrum) contributions were clearly seen in the spectra of connective tissue at 850, 940 and 1245 cm^{-1} (see the Raman spectrum of collagen, Figure 2.12) . Similar results were observed by De Jong *et al.* (2002).

Glycogen bands seen in spectra of cervical epithelial cells, were also observed by Choo-Smith *et al.* (2001) in cultured microcolonies. Levels of glycogen were seen to increase as the bacteria aged. The same parallel can be drawn with the epithelial cells recorded in Figure 6.5. As normal epithelial cells mature, glycogen levels increase as glucose is stored by the cell in the form of glycogen. Glycogen concentrations were at their highest in the superficial epithelial cells, which are the most mature cells.

Although great improvements were made in levels of wax removal from both ectocervical epithelial cells and connective tissue samples, wax contributions were surprisingly strong in the spectra of columnar cells. The reasons for this remain unclear. Absence of glycogen

from columnar cells has been seen previously in studies using FTIR spectroscopy (Wong *et al.* 2002) and results seen in the present study confirm these findings. Columnar cells can be identified by sharp bands seen in spectra at 717 and 1601 cm^{-1}.

The degree of spot to spot variation within a single sample was examined in normal epithelium and invasive carcinoma and was found to be minimal. In the spectra of normal epithelial cells, the main source of variation came from changes in overall intensity. All bands were present in spectra from each point and the shape of each of the overall spectral fingerprint was common to each point examined. The only changes to relative intensity in spectra from different samples came from small quantities of residual wax seen at 1063, 1296 and 1441 cm^{-1}, as previously discussed in Chapter 4. A slightly greater degree of variation was seen between spots in invasive carcinoma. Slight changes in relative intensity were seen at 726, 781, 1238 and 1576 cm^{-1}. These peaks are associated with nucleic acids. Therefore, there appears to be changes in concentration of nucleic acid in different spots, although at no time do these markers entirely disappear.

Patient to patient variability was examined in Figure 6.8 (normal epithelium) and Figure 6.11 (invasive carcinoma). Again the results are encouraging, apart from small changes in levels of residual wax and changes in the intensity of the peak at 1028cm^{-1} in the spectra recorded from invasive carcinoma.

When normal epithelial spectra were compared to spectra recorded from invasive carcinoma regions, consistent and very distinct differences between the spectra were obvious. The strong glycogen bands associated with normal epithelium (482, 849, 938,

1082 and 1336 cm^{-1}) were absent or significantly reduced in the spectra of invasive carcinoma. New bands appeared at 724, 779, 829, 852, 1098, 1240, 1366, 1484 shoulder and 1578 cm^{-1}.

To better understand the features of and the differences between normal tissue and invasive carcinoma, spectra were recorded from potential molecular constituents that were most likely to contribute to the spectral signature of the tissue. Spectra were recorded from glucose, glycogen, RNA and DNA. Figure 6.14(A) showed the spectrum recorded from glycogen, compared to the spectrum of normal epithelial tissue (Figure 6.14(C)). It is clear that glycogen features very strongly in the spectrum of epithelial cells. In order to rule out similar carbohydrates, Raman spectra were recorded from glucose (Figure 6.14 (B)). The presence of the peaks at 406 and 542 cm^{-1} present in spectra of glucose and absent in both glycogen and tissue spectra, confirms that it is indeed glycogen present in the tissue. Furthermore, Raman spectra were recorded from endocervical columnar cells, known to be glycogen free (Wong *et al.* 2002). The absence of glycogen features in these spectra further confirms the above hypothesis. In order to confirm that the changes seen in carcinoma tissue are the result of increased nucleic acid concentration, Raman spectra were recorded from both RNA and DNA and compared to the spectrum of invasive carcinoma. RNA and DNA can be differentiated by the double peak seen in RNA spectra at 809 cm^{-1}, which is not present in DNA spectra. This peak is not evident in the carcinoma spectra recorded and so it may be deduced that the increased levels of nucleic acid are a result of increased levels of DNA in the tissue. The increase in the Amide I band observed in the spectra of carcinoma tissue can also be attributed to DNA levels. Kneipp *et al.* (2003a) identified

DNA Raman contributions at 728, 782 and 1575 cm^{-1}, which are also observed in the DNA spectra recorded in the present study.

The bands at 829 and 852cm^{-1}, associated with O-P-O stretching and ring breathing of tyrosine respectively (Notingher *et al.* 2003), may be present in the tissue spectra of normal epithelium masked by the strong glycogen contributions.

In general the results described compare well with published data, but some differences were observed. Mahadevan-Jansen *et al.* (1998) recorded spectra from snap frozen biopsies of SIL, using 789nm excitation. It was found that peaks attributed to collagen vibrations (1656, 1070 cm^{-1}) decreased in intensity in spectra of tissues with SIL, while peaks attributed to vibrations of phospholipids, DNA and glucose 1-phosphate (1454, 1330, 978 cm^{-1}) increased in intensity in spectra of tissues with SIL. The study attributed two large peaks (1070, 1656cm^{-1}) to collagen modes. This assignment was based on recording a spectrum of pure collagen. However, the large peak seen at 1070 cm^{-1} was not reported in other studies where collagen spectra were recorded (De Jong *et al.* 2002). It may be possible that the decrease in intensity of the band at 1070cm^{-1} could be the result of a decrease in concentrations of glycogen (as seen in this study at 1082cm^{-1}), as opposed to collagen. The second peak to decrease in intensity (1656cm^{-1}) is the Amide I band, which is present in spectra of many biological macromolecules and therefore cannot be exclusively assigned to collagen. The decrease in the intensity of the Amide I band was not seen in the present study, where the Amide I band actually increased in intensity. Mahadevan-Jansen *et al.* (1998) also reported increases in the intensity of the bands at 1454, 1330 and 978 cm^{-1}.

Again these changes were not seen in the present study, although the conclusion that DNA increases with SIL is supported by the data presented in this Chapter.

Yazdi et al. (1998) investigated resonance Raman spectroscopy (257nm excitation) of normal and malignant cultured cervical cells in PBS suspension. Cell spectra obtained closely resembled that of DNA with peaks around 1330, 1480 and 1580 cm^{-1}, due to the vibrations of the nucleotide bases. The ratios of the intensities of the Raman spectral peaks 1480/1614 cm^{-1} and 1480/1540 cm^{-1}, which are sensitive to the concentration of nucleic acids relative to cell proteins, were found to be higher in malignant cells than in normal cells. Resonance Raman at 257 nm results in very strong nucleic acid contributions. In the present study DNA peaks were identified at 1335, 1484 and 1578cm^{-1}, which are very similar to those seen by Yazdi *et al.*(1998) and increases in intensity were also seen in the bands at 1484 and 1578cm^{-1}. Although protein peaks identified in the present study would be best assigned as 1447cm^{-1} (CH_2 scissoring) and 1660cm^{-1} (Amide I) as opposed to 1614 and 1540 cm^{-1} used by Yazdi *et al.*, the increase in DNA concentration would result in the same increased intensity ratio in carcinoma cell spectra.

Utzinger *et al.* (2001) evaluated the potential of near-infrared (789 nm excitation) Raman spectroscopy for *in-vivo* detection of cervical precancers. As the diagnosis progressed from normal to inflammation to metaplasia and through squamous dysplasia, the intensity of the Raman peak at 1454 cm^{-1} increased (assigned as phospholipids and collagen modes). The intensity of the Raman peaks at 1330 and 1656cm^{-1} was greater for abnormal samples than normal samples. The ratio of intensities 1454/1656 cm^{-1} was greater on average for squamous dysplasia samples than all other tissue types, while the ratio of intensities

1330/1454 cm^{-1} was lower on average for samples with squamous dysplasia than all other tissue types. Again, increases in the intensity of the band at 1454cm^{-1} were not seen in the present study. The increase in the intensity of the 1330 cm^{-1} band was not seen, although a decrease in the intensity of the 1336cm^{-1} band was observed as a result of reduction in glycogen content. The increase in intensity seen by Utzinger *et al.* in the band at 1656cm^{-1}, the Amide I band, was also observed in the present study. The ratio of intensities 1454/1656 cm^{-1} would be less in spectra of carcinoma tissue, due to the intensity increase in the Amide I band and not greater as seen by Utzinger *et al* (2001). The ratio of intensities 1330/1454 cm^{-1} would be higher due to a decrease in glycogen levels and not lower as seen by Utzinger *et al.*. However considerably more spectral features and chemical information were observed in the spectra from the present study than those obtained by Yazdi *et al.* This is probably due to a lower power (≈4mW) at the sample being used and the fact that spectra were being recorded from single cells in solution (PBS) by Yazdi *et al.*

Liu *et al.* (1992) observed a decrease in the intensity of the Amide I band in malignant tissue spectra, again not observed in the present study. However, interestingly, a band at ≈934 cm^{-1} was observed in spectra of normal and benign tissue, but not in spectra of malignant tissue. This band was unassigned but it is possible that the band could be glycogen vibrations and would therefore correlate with the present study. Liu *et al.* (1992) also observed an increase in the intensity of the band at 1262cm^{-1} in benign tissue spectra (similar to the band at 1240cm^{-1} observed in this study) which could be attributed to vibrations of nucleic acids, although this increase was not found in spectra of cancerous tissue. A peak which was not highlighted by Liu *et al.* but was present in the spectra

presented in the study, at around 1578cm^{-1}, appears to increase in intensity in both benign and malignant tissue spectra and would also confirm results seen in the present study.

The data presented in the present study did not confirm the findings of Mahadevan-Jansen *et al.* (*in vitro*) and Utzinger *et al.* (*in vivo*). Both studies examined precancerous tissue, such as intraepithelial lesions and squamous dysplasia using 789nm laser excitation. The present study focussed on invasive carcinoma, although neoplastic tissue was also examined with 514.5nm excitation. The spectra displayed in both studies differ somewhat from spectra obtained in the present study. The fact that samples were frozen in the *in vitro* studies should not affect the properties of the tissue, as seen in Chapter 4. A possible explanation for the differences may be that neither system was calibrated for spectral response. 789 nm excitation is close to the CCD detector cut off and far from the optimum maximum 500nm blaze grating quoted in the papers, which could result in a distorted spectrum. Measurement conditions at 514.5nm are close to the optimum conditions of the detector and the 600 nm blazed grating used here. This distortion might explain the Amide I band being much less pronounced in the Mahadevan-Jansen *et al.* study.

This study has resulted in very clear and consistent spectral markers of cancer. It has shown that good quality Raman spectra can be obtained using 514.5nm laser excitation (unlike Liu et al. 2005). The most marked changes in the spectra were in bands at 482, 849, 938, 1082 and 1336cm^{-1} associated with glycogen in normal tissue, which were absent in carcinoma spectra. Bands at 724, 779, 1098, 1240, 1366, 1484 and 1578cm^{-1} were found to increase in intensity in cancerous tissue spectra and have been attributed to increased concentration of nucleic acids. Bands were also seen to increase in intensity at 829, 852 and 1002cm^{-1} in

cancerous tissue. Chapter 7 describes the multivariate analysis of the spectra shown in this Chapter.

Chapter 7

Multivariate analysis of Raman spectral data

7.1 Introduction

With ten spectra recorded from fifty regions within the samples investigated, 500,000 wavenumber and 500,000 corresponding intensity values for the Raman spectra have been recorded. It is clear that a data reduction method and a multivariate analysis technique is required to deal with the 1 million data points generated in the Raman study.

Other Raman studies have employed a variety of different analysis methods including; **peak height ratios** (Huang *et al.* 2003, Lau *et al.* 2003, Knudsen *et al.* 2002, Caspers *et al.* 2001, Utzinger *et al.* 2001 and Yadzi *et al.* 1998), **curve fitting analysis** (Kaminaka *et al.* 2002), **classical least squares and partial least squares analysis** (de Jong *et al.* 2002, Shafer-Peltier *et al.* 2002a,b, Caspers *et al.* 2001, Buschman *et al.* 2001, Buschman *et al.* 2000), **principal components analysis in conjunction with K-means clustering analysis** (de Jong *et al.* 2003, Kneipp *et al.* 2003, de Jong *et al.* 2002, Nijssen *et al.* 2002, Koljenović *et al.* 2002 and Mahadaven-Jansen *et al.* 1998), **principal components analysis in conjunction with linear discriminant analysis** (Lau *et al.* 2005, Bakker Schut *et al.* 2002 and Stone *et al.* 2002) and **artificial neural networks** (Gniadecka *et al.* 1997).

Principal Components Analysis (PCA) is one of the most common multivariate techniques used in spectroscopy (Notingher *et al.* 2005). The technique reduces the dimensionality of the measurement matrix with the goal of representing the data using a smaller number of factors or principal components (PCs). Thus spectra can be grouped, for example according to prominence of specific spectral features. However, PCA is not the optimal method for

spectral discrimination, because the PCA factors are computed to maximise the variance of the measurement matrix without considering sample groups. Linear Discriminant Analysis (LDA) is a more powerful method for discrimination of sample groups (Notingher *et al.* 2005). The main disadvantage of LDA is that the number of variables used as an input for an LDA model, has to be smaller than the total number of samples. Therefore entire Raman spectra cannot be analysed directly because the number of wavelength positions is usually larger than the number of samples analysed (Notingher *et al.* 2005). Hence data reduction by PCA is normally applied to the dataset. LDA is then used as a further refinement of the analysis procedure. PCA identifies differences between data sets and LDA maximizes these differences so as to group them. PCA in conjunction with LDA has therefore been used to analyse the data in the present study.

7.2 Multivariate Analysis

Chemometrics has been agreed as being "how to get chemically relevant information out of measured chemical data, how to represent and display this information and how to get such information into data" (Wold *et al.* 1988). It has found great success in applications where pattern recognition, classification and discriminant analysis are required. When faced by large sets of data, multivariate techniques can quickly reduce and simplify the problem. However there is a need for prudence in applying such techniques, as bad data cannot be changed into good data. In other words if the spectra recorded do not contain useful information, no amount of statistical analysis will change that fact.

7.2.1 Principal Components Analysis

Principal Components Analysis is a method that is mainly used for data reduction as it can maintain the total variance within the data set while reducing the number of variables used. PCA is an unsupervised method that changes the original data set into a relatively smaller and simpler one that can still adequately give the same amount of information, by removing redundant variables. Of course some information is lost but the improvement in ease of use of the information compensates for this.

Figure 7.1 shows a spread of experimental points in a two dimensional plane expressed in terms of the experimental measurables X and Y. Assuming that the origin for these axes is where the mean centre would be, it can be seen that the axes cover all the points. However using a different set of axes X' and Y' could reduce the density of the points per axis unit. These axes would still be orthogonal and would use the same origin (Figure 7.1). Therefore the "spread" or variance of points would be increased. This implies that the data can be better expressed in terms of the parameters X' and Y' which are combinations of the measured features X and Y. For example, the analysis methods based on peak height ratios (Huang *et al.* 2003, Lau *et al.* 2003, Knudsen *et al.* 2002, Caspers *et al.* 2001, Utzinger *et al.* 2001 and Yadzi *et al.* 1998) group the spectra in terms of ratios of identified spectral markers rather than the heights of the individual markers.

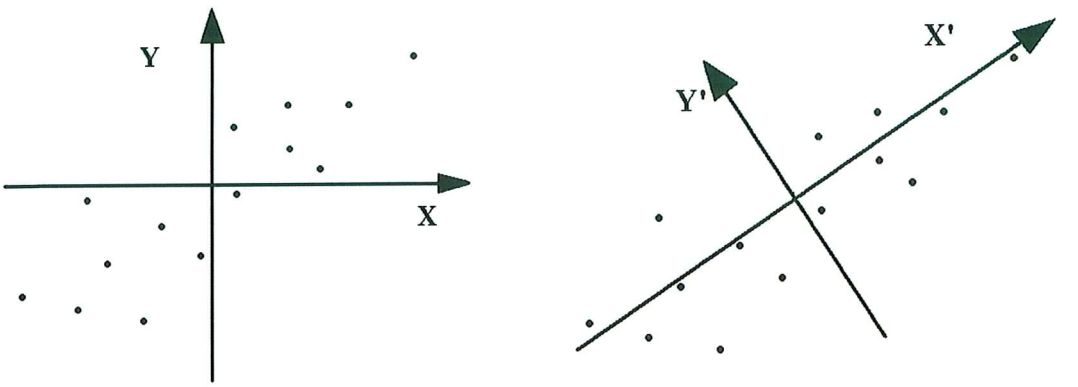


Figure 7.1 Rotation of orthogonal axes to align with greatest variance in data points

To reduce the number of descriptors or variables, data points must be transformed onto a new axis. To do this data points are projected onto a new axis, whose orientation maximises the variance of the data. Figure 7.2 illustrates how a 2-D system can be reduced to a 1-D system by orthogonal projection of all data points onto the new axis, X'.

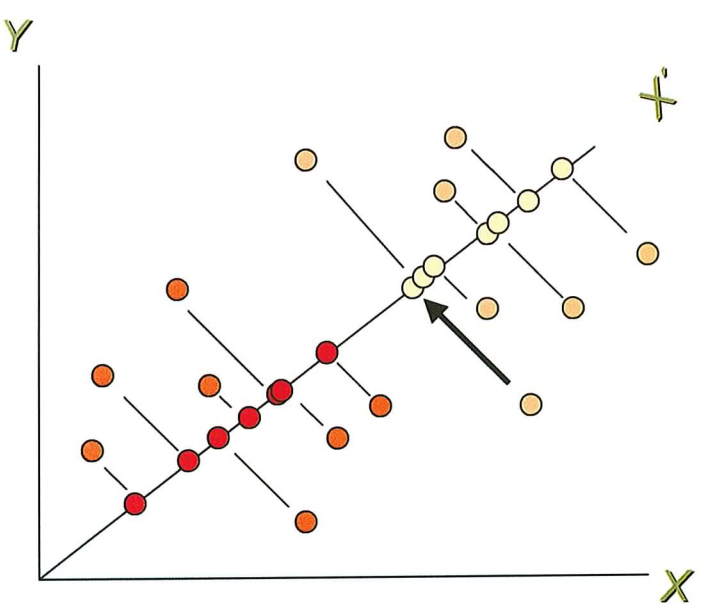


Figure 7.2 Data reduction from a 2-D system to a 1-D system

From the original data set the first principal component is calculated and this contains most of the information about the variance of the data set. The second principal component, being orthogonal to the first, then contains information on the remaining variances and so on for subsequent principal components. The original data matrix is a product of a scores matrix and transpose of the corresponding loading matrix (Figure 7.3).

The general equation is

$$X = TL^T \qquad \textit{Equation 7.1}$$

where X is the original data matrix with n rows (objects) and p columns (features), T is the scores matrix with n rows and d columns (Principal components) and L^T is the transpose of the loading matrix with p rows and d columns. Simply multiplying both sides by the inverse of L gives T, where the first set of principal components make up the first column and the second set of components make up the second column and so on. These components are graphically represented as PC1, PC2 *etc.* The characterisation of a sample is usually dominated by the principal component and the relative importance of subsequent PCs decreases with their order.

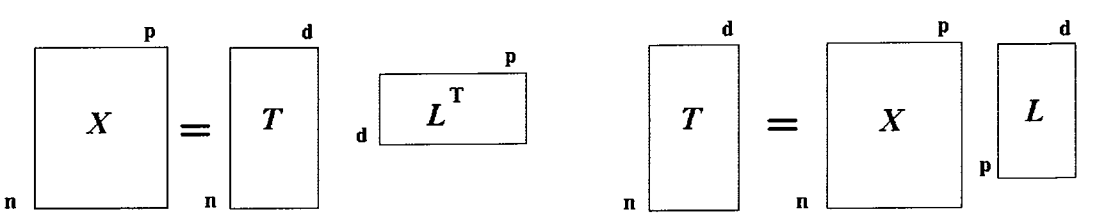


Figure 7.3 PCA explained though matrices

Essentially this method generates eigenvectors for the defined eigenvalues. These eigenvectors then relate the amount of information that the eigenvalue contributes to the overall variance. Those with the greatest eigenvectors are therefore the most important.

As applied to spectral analysis, multicomponent analysis generates principal components which represent combinations of all spectral features, weighted or "loaded" according to their relative contributions.

7.2.2 Loading Plot

In a PC-model with two principal components, it is desirable to find out which features are responsible for the patterns seen among the observations (it is of interest to know which features are influential in determining the principal components and also how they are correlated). Such information is given by the principal components **loading plot**. These loading vectors display the relationships between all contributions in the data set (spectrum) at the same time. Features contributing to a similar degree are grouped together, that is, they are correlated. When the numerical value of one contribution increases or decreases, the numerical value of the other contribution has a tendency to change in the same way. When they are negatively ("inversely") correlated they are positioned on opposite sides of the plot origin, in diagonally opposed quadrants (Eriksson *et al.* 2001).

7.2.3 Derivative Spectra

In spectroscopy, differentiation of spectra is a widely used technique, referred to as derivative spectroscopy. Derivative methods have been used in analytical spectroscopy for three main purposes: (a) spectral discrimination, as a qualitative fingerprinting technique to accentuate small structural differences between nearly identical spectra; (b) spectral resolution enhancement, as a technique for increasing the apparent resolution of overlapping spectral bands in order to more easily determine the number of bands and their wavelengths; (c) quantitative analysis, as a technique for the correction for irrelevant background absorption and as a way to facilitate multicomponent analysis.

Spectra commonly consist of one or more spectral bands, usually of approximately Gaussian or Lorentzian shape and are often superimposed on a background (baseline) of variable shape. One of the best methods for removing baseline effects and enhancing spectral maxima and minima is the use of derivative spectra. In the 1st derivative spectrum, the baseline is slowly varying in comparison to the spectral features and can be removed and maxima and minima of curves can be pinpointed as they are zero in the derivative spectrum. The 2nd derivative is a measure of the change in the slope of the curve. In addition to removing the baseline offset, it is not affected by any linear "tilt" that may exist in the data and is therefore a very effective method for removing both the baseline offset and slope from a spectrum. It also identifies points of inflexion of the spectral features giving further information as to their shape. The main disadvantage of using derivative pre-processing is that the resulting spectra can be difficult to interpret.

7.2.4 Linear Discriminant Analysis (LDA)

For some types of analysis it might be necessary to distinguish between two different known populations. To this end it might be necessary to devise a rule or function that would allocate an object to a specific group. Such a function is known as a discriminant function and the type of analysis is known as Linear Discriminant Analysis. Therefore it is necessary to construct a classification rule which will allocate individuals to pre-assigned groups with the lowest error rate. Discriminant functions aim to maximise separation between groups of (available) individuals, while classification rules aim to minimise the misclassification rate over all possible (future) allocations. Probabilistic classification rules are based on the premise that a large number of individuals will in the future need to be classified and hence the classification rule should be chosen in such a way so as to minimise the expected consequences of mistakes made in this series of allocations (Wold *et al.* 1988).

A brief simplified mathematical explanation is now given:

In LDA the *Standard Distance* (D) gives a dimensionless measure of the separation between two groups (similar to *t*-statistic)

$$D = \frac{\left|\overline{X_1} - \overline{X_2}\right|}{s} \qquad \textit{Equation 7.2}$$

where $\overline{X}$ is the mean of groups 1 and 2 and s is the standard deviation. Any member of a group can be described using the expression shown in Equation 7.3:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 .. + \beta_\pi X_\pi \qquad \textit{Equation 7.3}$$

and will have its own coefficients β_0, β_1, β_2, β_3.. β_π.. Therefore each group will have a mean and standard deviation, implying that the standard distance will be dependant on the values of the coefficients β_0, β_1, β_2, β_3.. β_π. (Equation 7.4);

$$D(Z) = D(\beta_0, \beta_1, \beta_2, \beta_3 .. \beta_\pi) \qquad \textit{Equation 7.4}$$

and therefore the Multivariate Standard Distance, D_p, will be the greatest distance attainable between the two groups as shown in Equation 7.5.

$$D_p = \max D(\beta_0, \beta_1, \beta_2, \beta_3 .. \beta_\pi) \qquad \textit{Equation 7.5}$$

The linear combination of coefficients for which the maximum D(Z) is achieved is called the Discriminant Function. Using this function it will be possible to create a classification rule that would assign unknowns to pre-determined groups.

In a visual sense this means that if one looks at the two groups of data in Figure 7.4 from the traditional orthogonal X view (green arrow), there is a region of overlap where group membership would be shared. The same would be true for the corresponding Y view. However if a view from an alternative angle is taken, X' (yellow arrow), then a clear

separation is obtained. So essentially LDA aims to find the best angle to look at data so as to maximise the space between known groups (Duffy 2002).

When used in combination with Principal Components Analysis, LDA groups the samples with similar correlations of the principal components identifying them as being of similar origin. PCA performs a feature-led classification of data and LDA groups the classifications. Thus if a group of spectra have a similar correlation to the shape defined by PC1 and that defined by PC2, they are grouped together.

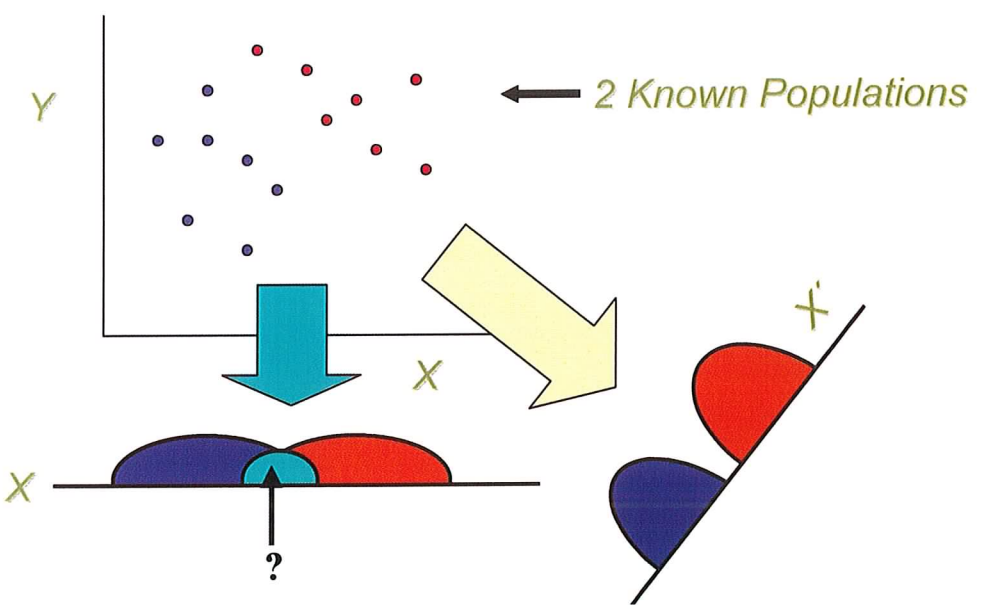


Figure 7.4 Visual interpretation of LDA

7.2.5 Leave-one-out cross-validation

Cross-validation is a method of estimating how well the model derived from the training data will perform on future as-yet-unseen data.

In cross-validation, the subsample of data that is used for the initial analysis is called the training data. In systems where models are autonomously constructed from the training data (as in this case), the process of constructing the model is called training. The subsamples of data that are used to validate the initial analysis (by acting as "blind" data) are called validation data, or test data.

As the name suggests, leave-one-out cross-validation involves using a single observation from the original sample as the validation data and the remaining observations as the training data. This is repeated such that each observation in the sample is used once as the validation data. This is the same as K-fold cross-validation where K is equal to the number of observations in the original sample (Brereton 2005).

7.2.6 Sensitivity and Specificity

Many studies quote sensitivity and specificity values associated with techniques for the detection of cancer. Sensitivity and specificity are the most widely used statistics used to describe a diagnostic test. The sensitivity refers to the probability of a positive test among patients with disease, while the specificity refers to the probability of a negative test among patients without disease. A single parameter is insufficient, as either 100% sensitivity or

100% specificity can be achieved by labelling all tests as positive or negative respectively. Sensitivity and specificity values can be calculated using the table and formulas shown below (Equations 7.6 and 7.7, in conjunction with Table 7.1);

$$Sensitivity = \frac{a}{a+c} \qquad \textit{Equation 7.6}$$

$$Specificity = \frac{d}{b+d} \qquad \textit{Equation 7.7}$$

	Number of patients with disease	Number of patients without disease
Test is positive	**a**	**b**
Test is negative	**c**	**d**

Table 7.1 Explanation of values given in Equations 7.6 and 7.7

For a group of patients diagnosed as either with or without disease, ***a*** represents the number of those patients with disease classified correctly, ***c*** represents the number of those patients with disease classified incorrectly, ***b*** represents the number of patients without disease, classified incorrectly and ***d*** represents the number of patients without disease classified correctly. These calculations can be carried out for all patient types.

Sensitivity and specificity values produced by other Raman studies, using various analysis methods and tissue samples, usually range from 70 -100%. Lau *et al.* (2005) achieved prediction sensitivities of 89%, 69% and 88% and specificities of 86%, 94% and 94% for normal tissue, carcinoma and squamous papilloma using multivariate techniques. Nijssen *et al.* (2002) achieved 100% sensitivity and 93% specificity discriminating basal cell carcinoma from its surrounding tissue using a logistic regression model. Stone *et al.* (2002) used multivariate methods (principal components fed linear discriminant models) to discriminate epithelial pre-cancers and cancers of the larynx, tonsil, oesophagus, stomach, bladder and prostate. Larynx samples, with squamous epithelial tissue, were separated into three distinct groups with sensitivities ranging from 86 to 90% and specificities from 87 to 95%. Bladder specimens, containing transitional epithelial tissue, were separated into five distinct groups with sensitivities of between 78 and 98% and specificities between 96 and 99%. Oesophagus tissue can contain both squamous and columnar cell carcinomas. A three group model discriminated the columnar cell pathological groups with sensitivities of 84-97% and specificities of 93-99% and an eight group model combining both columnar and squamous tissues in the oesophagus was able to discriminate pathologies with sensitivities of 73-100% and specificities of 92-100%. Koljenović *et al.* (2002) discriminated vital tumour from necrotic tissue in human glioblastoma sample with an accuracy of 100%, while Notingher *et al.* (2005) also used multivariate analysis to discriminate between healthy and tumour cells with a prediction accuracy of 95%.

With respect to cervical tissue, Mahadaven-Jansen *et al.* (1998) differentiated precancers from other tissues using Raman spectroscopy with an average sensitivity and specificity of

88% and 92% repsectively. Ratios of un-normalised intensities differentiated precancers from other tissues with a sensitivity and specificity of 82% and 88% respectively and high-grade from low-grade lesions with both sensitivity and specificity of 100%. Using multivariate methods, intensities at eight frequencies were used to differentiate precancers from all other tissues with a sensitivity and specificity of 82% and 92% respectively in an unbiased test.

This Chapter aims to evaluate a number of different methods of data preparation for PCA and LDA, including specific wavenumber selection, averaging spectra over a $10cm^{-1}$ range, spectra normalisation, as well as calculating 1^{st} and 2^{nd} derivatives of spectra. Sensitivity and specificity values obtained will be compared with published data.

7.3 Data Analysis

(i) Pre-processing

All spectra (unfiltered) presented in Chapter 6 were subjected to spurious peak ("cosmic ray") removal and baseline correction using a common baseline, in Labspec (v. 4.02 Jobin Yvon), before being exported in ascii format to Microsoft Excel.

(ii) Matrix Manipulation

Basic matrix manipulations, data reduction and calculation of derivative spectra were carried out in Microsoft Excel Professional 2003 (v. 11.0), before being exported into Minitab to perform PCA and LDA.

(iii) PCA and LDA Analysis

Minitab Release 14.1 Statistical Software Analysis Programme was used to produce Loading plots, PCA plots and LDA plots, as well as to carry out leave-one-out cross validation.

7.4 Results

7.4.1 Intrapatient and Interpatient Variability

Before applying discriminant analysis to the datasets, the intra and interpatient variability was first examined. Figure 7.5 shows the normalised, average spectra of the 3 tissue classes examined – both intra and interpatient, along with the standard deviation associated with each class. The glass peak at 560cm^{-1}, seen previously (Figure 2.17), contributes a large degree of variation to all samples. The spectra of normal tissue samples appear to be producing the most variation among the three tissue types. In order to visualise this variation as it applies to intra and interpatient variability, a plot of the PCs of each the normal samples was generated (Figure 7.6).

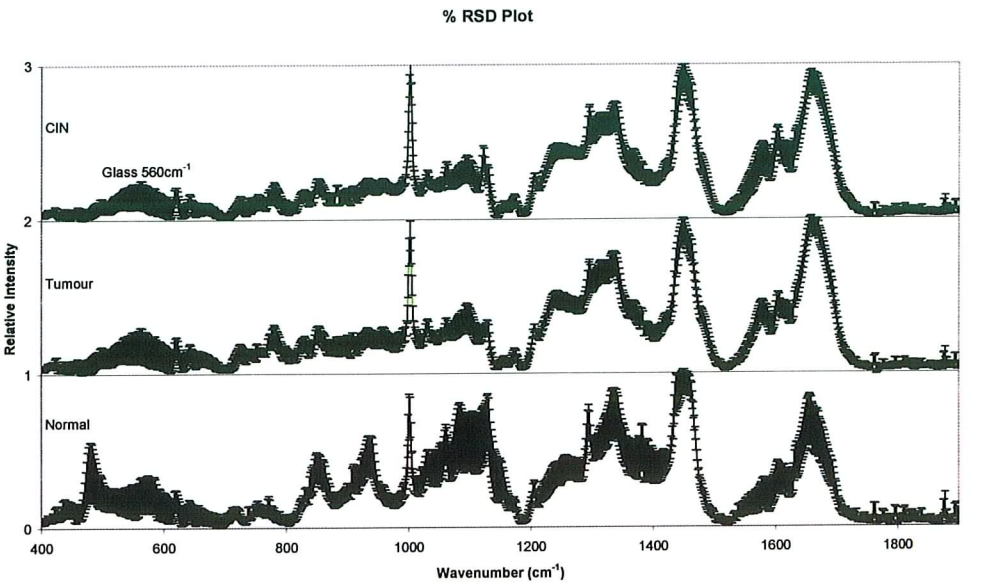


Figure 7.5 Average spectra of normal, invasive carcinoma and CIN samples, with associated standard deviation

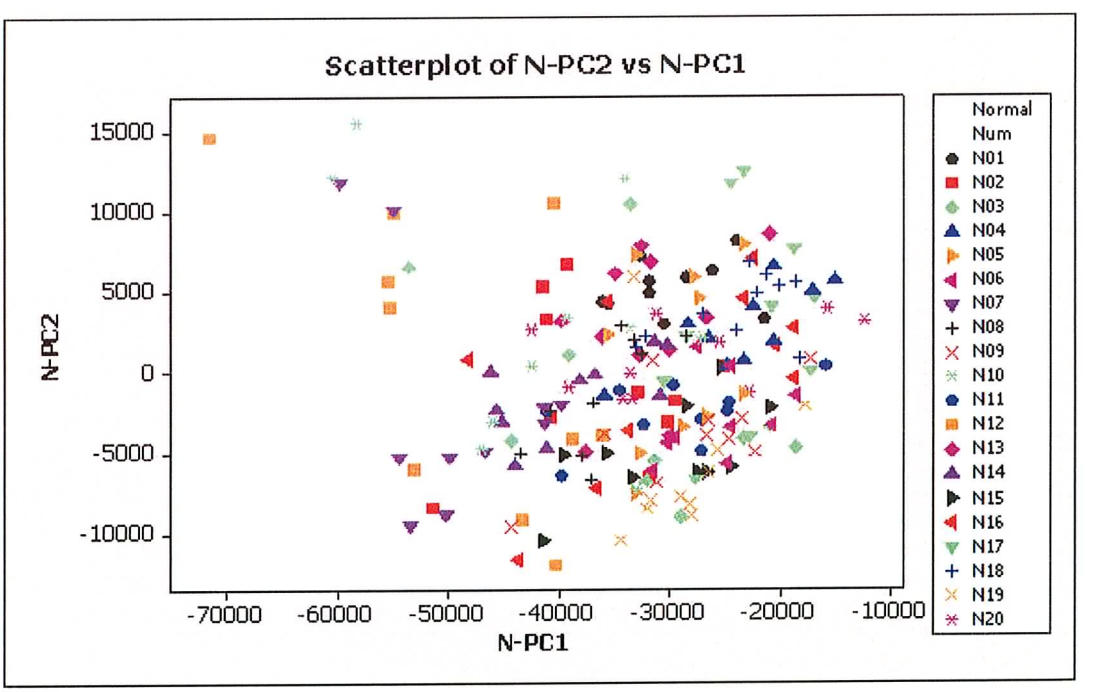


Figure 7.6 Scatterplot of intra and inter patient variability

Figure 7.6 shows the cluster analysis PCs for each of the normal samples (N01-N20). Each of these interpatient samples is represented by ten independent intrapatient measurements. It is clear from the plot that there is a significant distribution of points, both intra and interpatient. As suggested by the spectra presented in Chapter 6, the intrapatient variability is less than the interpatient variability but the fact that no single sample is clustered is a positive result and there is a large overlap between the different sets of measurements. This implies that the spectra obtained contain relatively generic information without undue patient specific properties.

7.4.2 Specific Ranges

In order to evaluate the most effective method of data analysis, a step by step approach was taken to different methods of data processing prior to PCA and LDA. The first method examined involved normalising spectra and inputting a number of specific wavenumber regions identified as possible markers of both normal and diseased tissue as outlined in the previous Chapter. These peaks included normal markers (482, 849, 938, 1082 and 1336 cm^{-1}, attributable to glycogen) and abnormal markers (724, 779, 1098, 1336 and 1578 cm^{-1}, attributable to nucleic acids). Wavenumber regions of $4cm^{-1}$ either side of the peak values identified were included. Figure 7.7 shows the regions selected, in relation to the spectra of normal and invasive carcinoma tissue. This data was then input into PCA where the first 4 principal components were calculated (accounting for over 95% of the variance). The resulting scatter plot is shown in Figure 7.8.

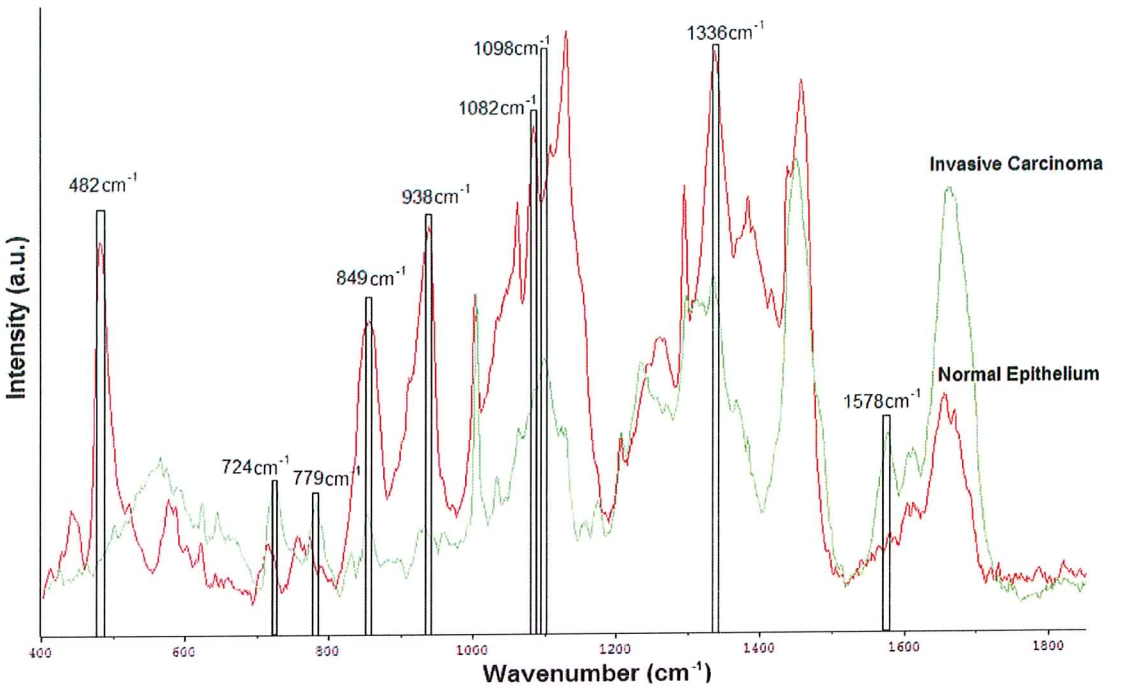


Figure 7.7 Regions selected for specific analysis

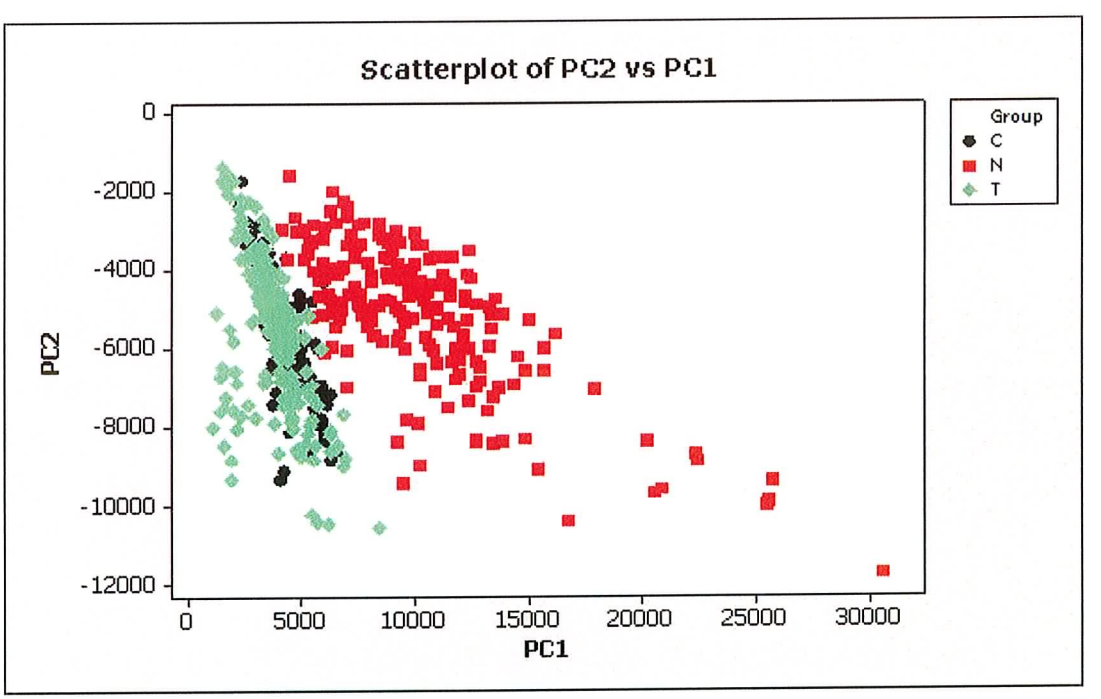


Figure 7.8 PC scatterplot of normalised specific wavenumber regions

(where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

It is clear that inputting these 9 wavenumber regions alone results in relatively good separation between normal and abnormal tissue, although separation between CIN and invasive carcinoma tissue is poor. Figure 7.9 is the loading plot associated with Figure 7.8. This plot visually demonstrates the criteria used to generate the scatterplot. The two plots can in fact be superimposed. The further apart the point on the loading plot the greater the difference between the spectra based on that point It is clear that the PCs used to identify abnormal tissue were 724, 779 and 1578 cm^{-1} (lower left quadrant) and the peaks used to identify normal tissue were 482, 849, 938, 1082 and 1098 cm^{-1} (upper right quadrant). The peak at 1336 cm^{-1} (identified as a normal marker in the previous Chapter), was found to be inadequate to differentiate between the two classes of tissue, as it is approximately equidistant from both quadrants. This peak (assigned to CH_2 deformation) is prominent in the spectra of both normal and abnormal tissue (Figure 6.13). The difference in the shape of the band between normal and abnormal is apparent to the eye when the data is being examined empirically. However by selecting only a narrow range and normalising the data this difference may not be as apparent (Figure 7.7). The classification of the 1098 cm^{-1} peak as a normal marker is also due to a similar phenomenon. Although by eye, it is clearly a peak present in spectra of abnormal tissue (Figure 7.7), spectra of normal tissue also have high intensity values in the region and hence the distinctiveness of the abnormal marker is "masked".

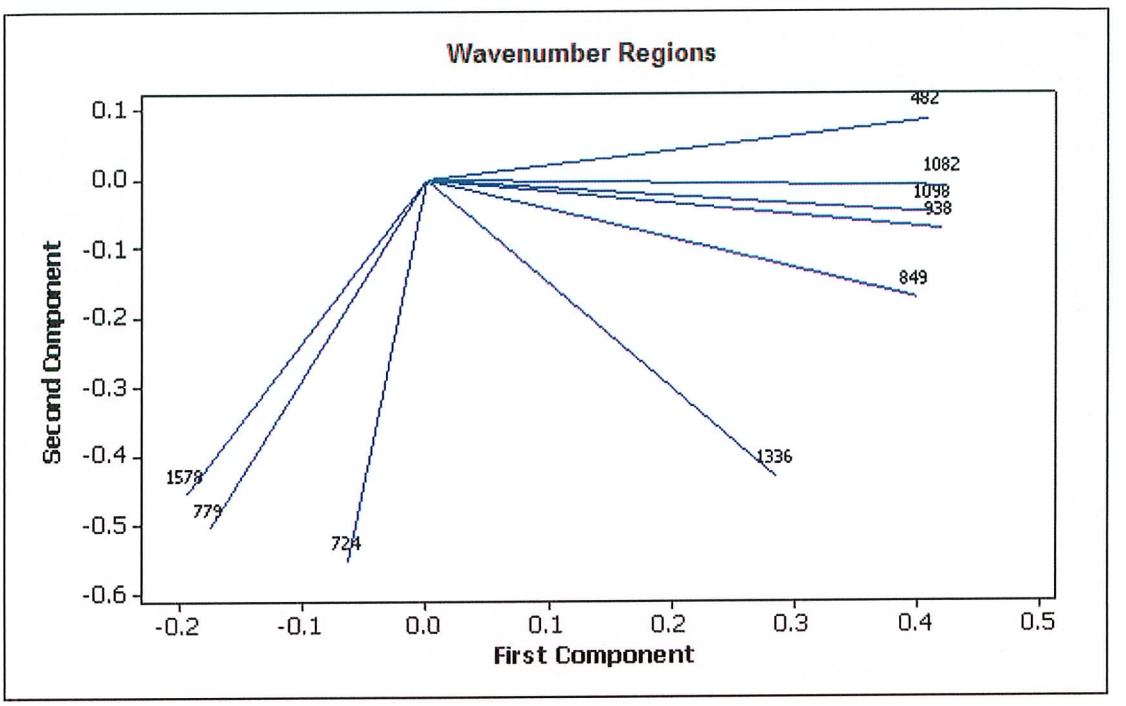


Figure 7.9 Loading plot of normalised specific wavenumber regions

By applying these principal components values to the LDA model, the differences between the PCs of the three tissue classes can be maximised. Figure 7.10 shows the scatterplot associated with the PCs above and Table 7.2 shows the classification results using cross validation.

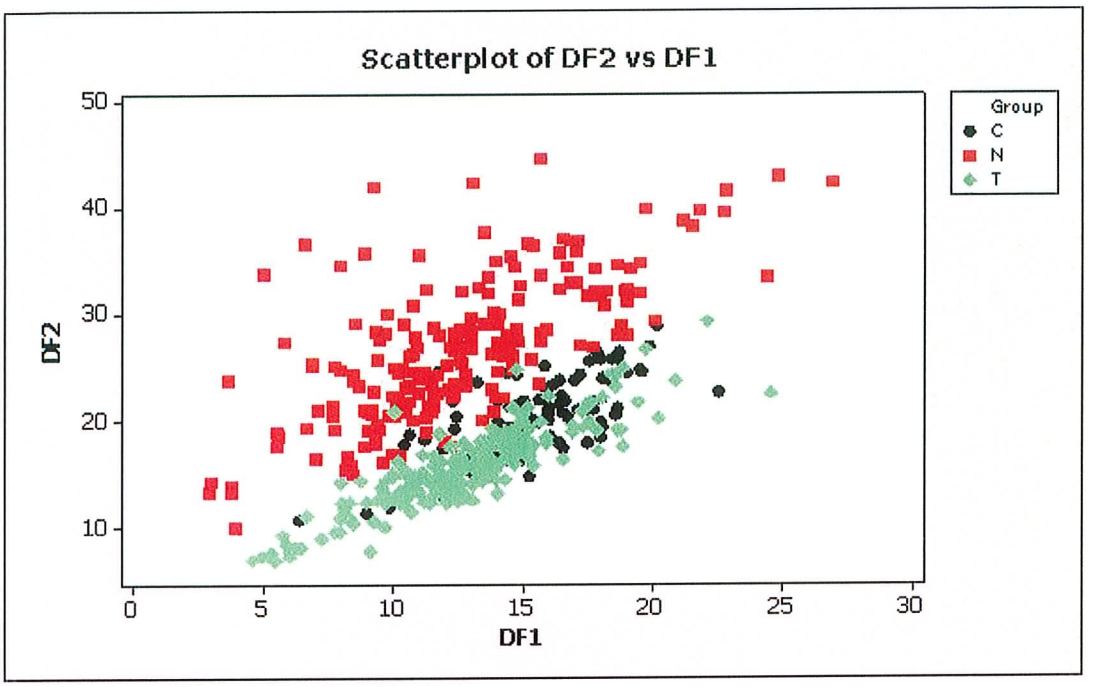


Figure 7.10 LDA plot of PCs generated from normalised specific wavenumber regions

(where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	75	15	34
N	3	184	1
T	22	1	163
Total N	100	200	198
N correct	75	184	163
Proportion	0.750	0.920	0.823

N = 498
N Correct = 422
Proportion Correct = 0.847

Table 7.2 Results of classification for normalised data at specific wavenumbers

Using LDA, with leave-one-out cross validation, 422 of the 498, spectra were correctly classified. 16 normal samples were assigned as abnormal (15C and 1T) and more problematic was the assignment of 4 abnormal samples (3 of C and 1 of T) as normal. Differentiation between both abnormal samples (CIN and invasive carcinoma) is poor with a total of 56 misclassifications.

In order to see whether or not taking the derivative spectra would improve classification, the 1st derivative of the normalised data was calculated. As the derivative spectra no longer contain peak maxima positions, the associated loading plot also no longer contains information about diagnostic peak positions. The previously outlined regions were then inputted into the PCA analysis (Figure 7.11) and the PCs applied to the LDA model (Figure 7.12)

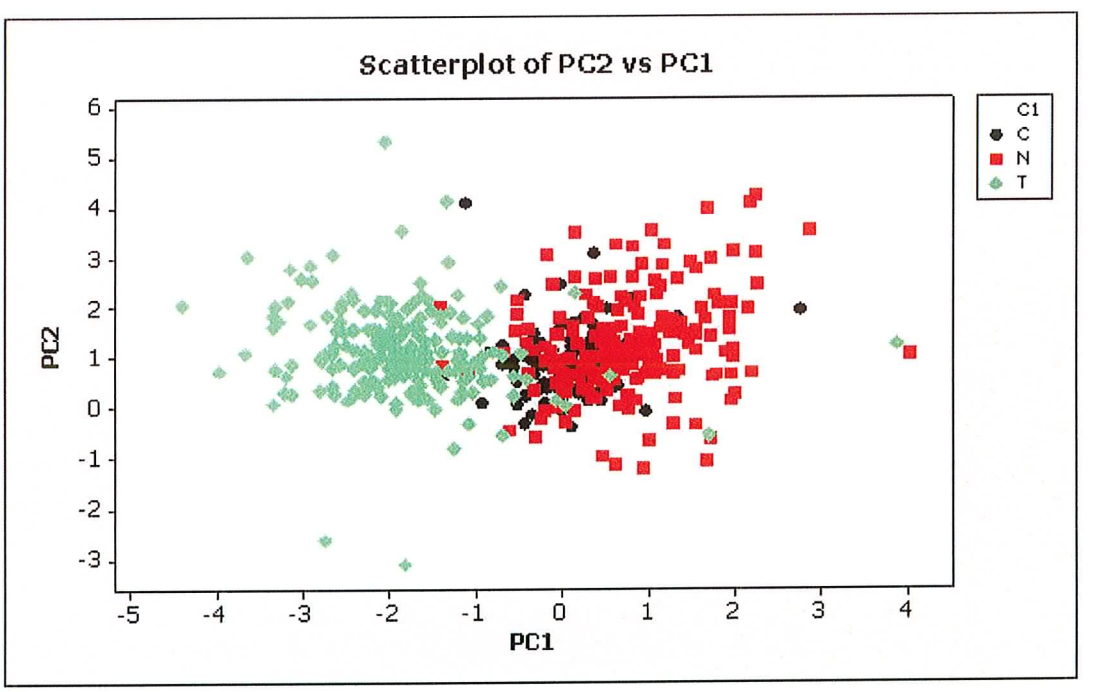


Figure 7.11 PC scatterplot of specific wavenumber regions of 1st derivative spectrum (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

The 1st derivative spectra also produce separation between the groups. Although in this two-dimensional plot the CIN tissue appears to be grouped with the normal tissue, there is separation when viewed in three-dimensions. When discriminant analysis is applied, the number of correct classifications increased to 458 (Table 7.3), compared with 422 in the undifferentiated spectra. It resulted in better separation/discrimination of both CIN and invasive carcinoma samples (with only 13 abnormal samples assigned to the wrong abnormal class (4 of C and 9 of T)). However taking the 1st derivative has resulted in 11 abnormal samples being classified as normal.

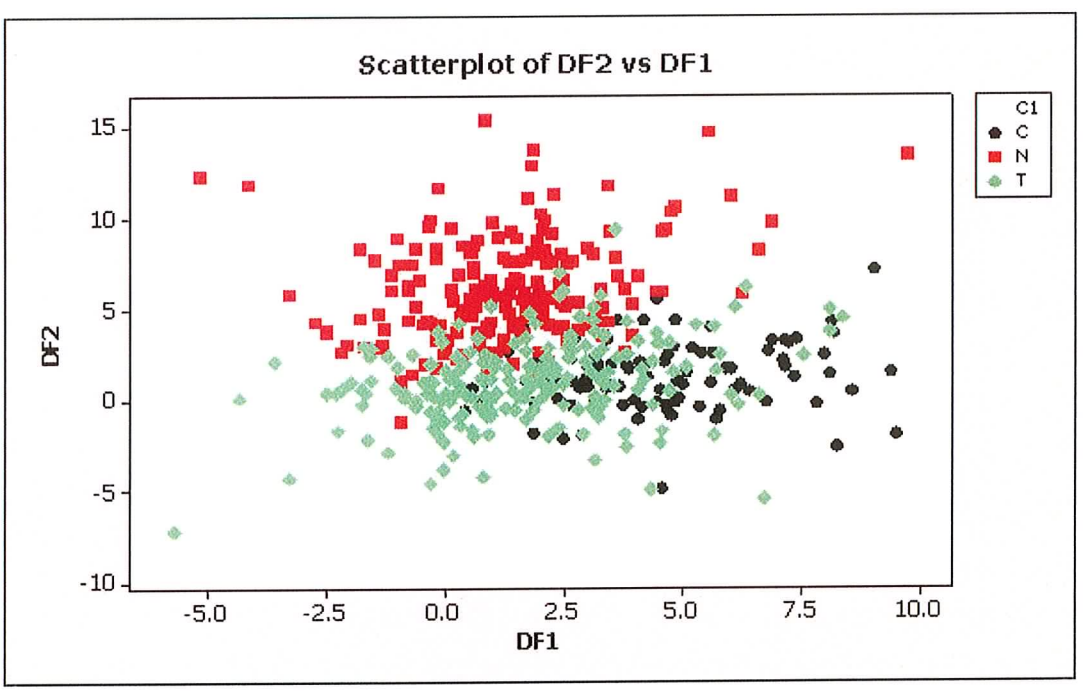


Figure 7.12 LDA plot of PCs generated from specific wavenumber regions of 1st derivative spectrum (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	90	12	9
N	6	184	5
T	4	4	184
Total N	100	200	198
N correct	90	184	184
Proportion	0.900	0.920	0.929

N = 498
N Correct = 458
Proportion Correct = 0.920

Table 7.3 Results of classification of specific wavenumber regions of 1st derivative spectra

The second derivative spectra were then calculated and again the regions selected previously were used as inputs into PCA analysis (Figure 7.13). It is clear that there is a much greater distribution of points with the PCs of the second derivative. This is most likely due to the fact of an over sampling effect by taking the second derivative. The first derivative identifies the positioning of the maxima and minima of a curve, which for an idealised (e.g. Lorentzian) Raman feature are spaced by the order of $10cm^{-1}$. The second derivative identifies points of inflexion which for such an idealised feature are of the order of $5cm^{-1}$. The spacing between the pixels in this wavelength range and at this resolution is $1.6cm^{-1}$. Therefore an idealised second derivative curve describing two maxima, a minimum and four zero points is defined by only ~ 10-20 data points. It is clear therefore that the integrity of the second derivative cannot be maintained by so few sampling points.

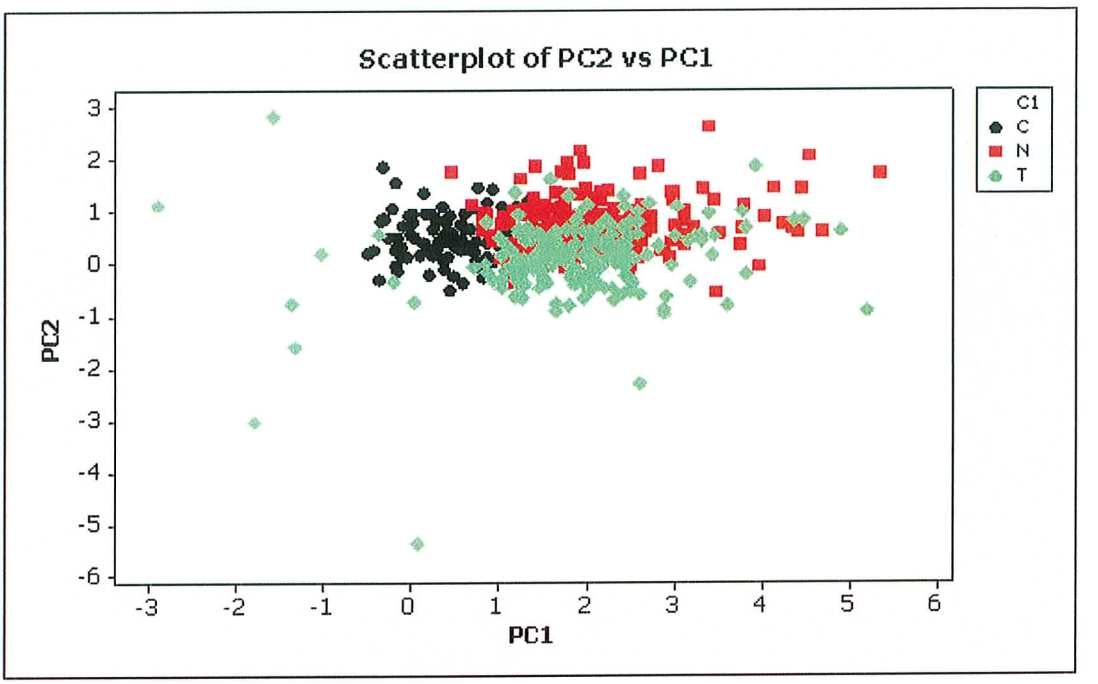


Figure 7.13 PC scatterplot of specific wavenumber regions of 2nd derivative spectra

(where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

When the principal components were applied to the LDA model (Figure 7.14), the number of correct classifications reduced to 370 correct out of 498 (Table 7.4). This is a marked reduction in classification performance and suggests that taking the second derivative of the Raman spectra does not enhance the differences between the spectra, but in fact contributes to a degradation of the spectral features, as outlined above.

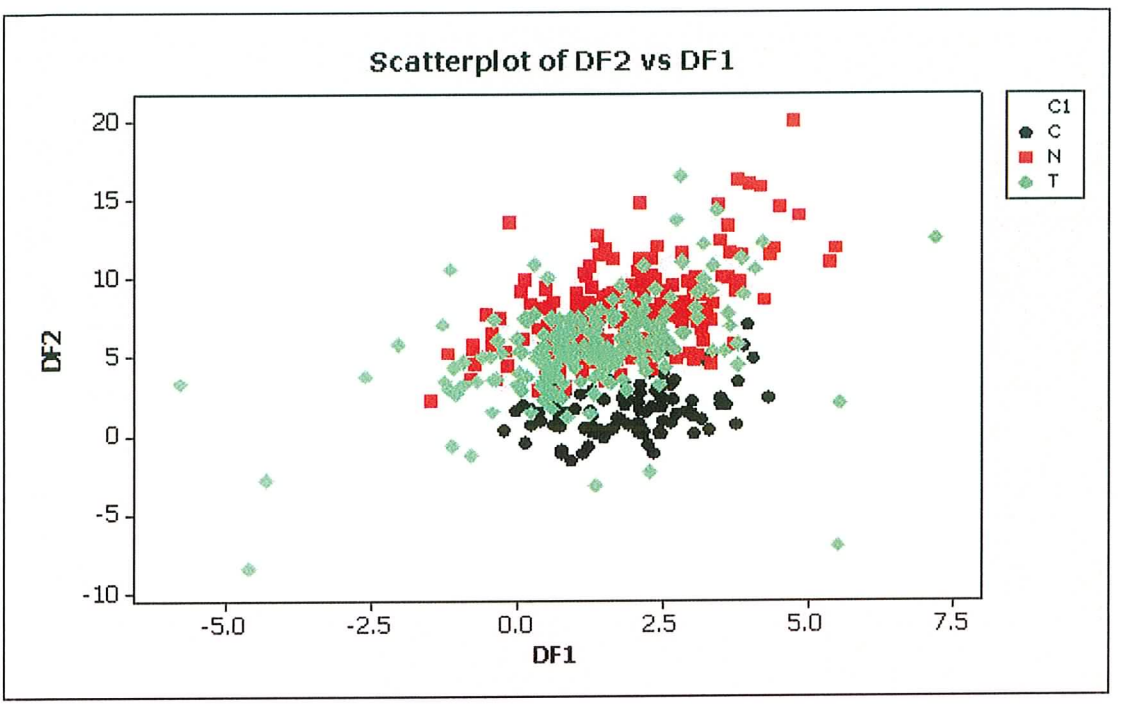


Figure 7.14 LDA plot of PCs generated from specific wavenumber regions of 2nd derivative spectra (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	91	12	19
N	3	141	41
T	6	47	138
Total N	100	200	198
N correct	91	141	138
Proportion	0.910	0.705	0.697

N = 498
N Correct = 370
Proportion Correct = 0.743

Table 7.4 Results of classification of specific wavenumber regions of 2nd derivative spectra

The next method of analysis involved applying the entire raw spectra to the PCA. This method with the vast amount of data resulted in the software package and computer slowing down dramatically, before crashing. It was decided therefore, to take a 10 point average in order to reduce the number of variables to be computed.

7.4.3 Entire Range - 10 Point Average

A ten point average was calculated for each spectrum, resulting in 1,500 wavenumber variables being reduced to 150 (R1-R150). Although this appears to be a dramatic reduction in the level of data, a large proportion of the spectral integrity is retained (Figure 7.15). The reduced data sets were applied to the PCA analysis. Figures 7.16 and 7.17 show the resulting scatter and loading plots respectively for the PCA analysis.

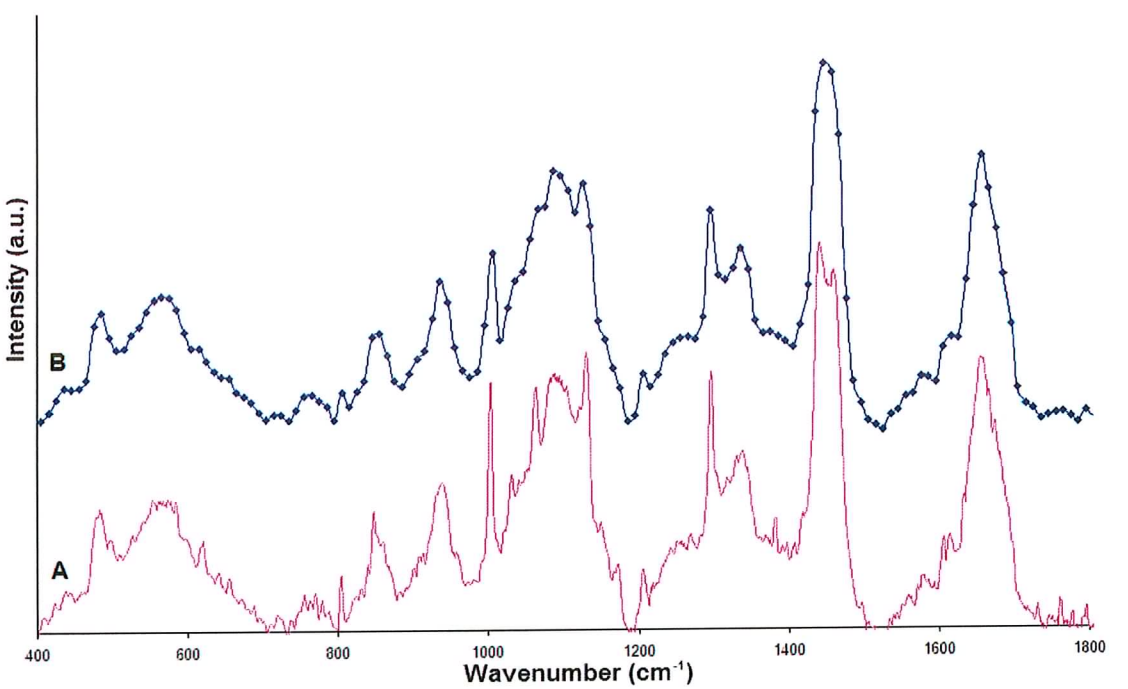


Figure 7.15 (A) Raw Raman spectrum and (B) 10cm^{-1} averaged Raman spectrum of normal tissue

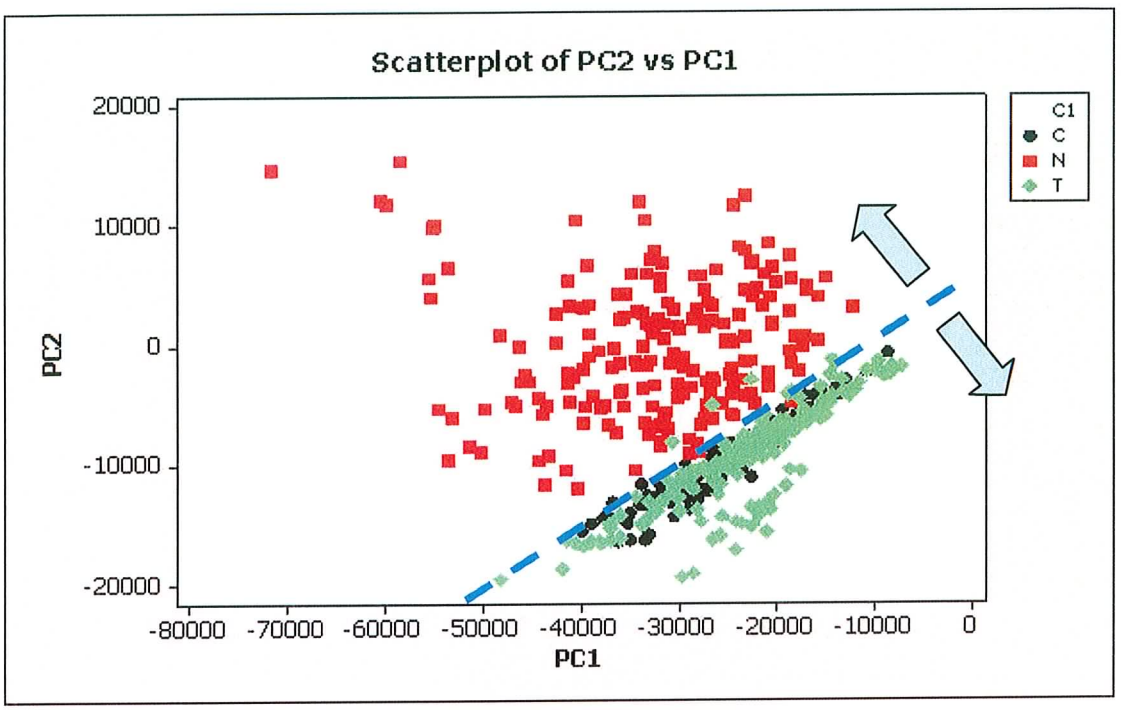


Figure 7.16 PCs scatterplot for 10 point average over entire spectral range

(where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

From the scatterplot it is clear that normal and abnormal tissues are separated, shown by the dashed blue axis. Normal samples have been grouped above this line and abnormal samples below (indicated by the arrows). Examining the loading plot (Figure 7.17) used to generate the scatterplot some of the empirically suggested markers of normal and abnormal tissue have been grouped together along a "10 o'clock" and "6 o'clock" separation. These have been marked as A (normal) and B (abnormal) on the diagram. Markers for normal (Fig. 7.17A) include R8 – R10 (479 cm^{-1}), R44 – R46 (849 cm^{-1}) and R54 – R55 (935 cm^{-1}), while some known markers for abnormal tissue (Fig. 7.17B) include R33 – R34 (725cm^{-1}), R38 – R40 (779 cm^{-1}) and R116 – R119 (1579 cm^{-1}). Re-examining Figure 7.16 it is clear

that the normal and abnormal points are separated by this "10 o'clock"/ "6 o'clock" criterion.

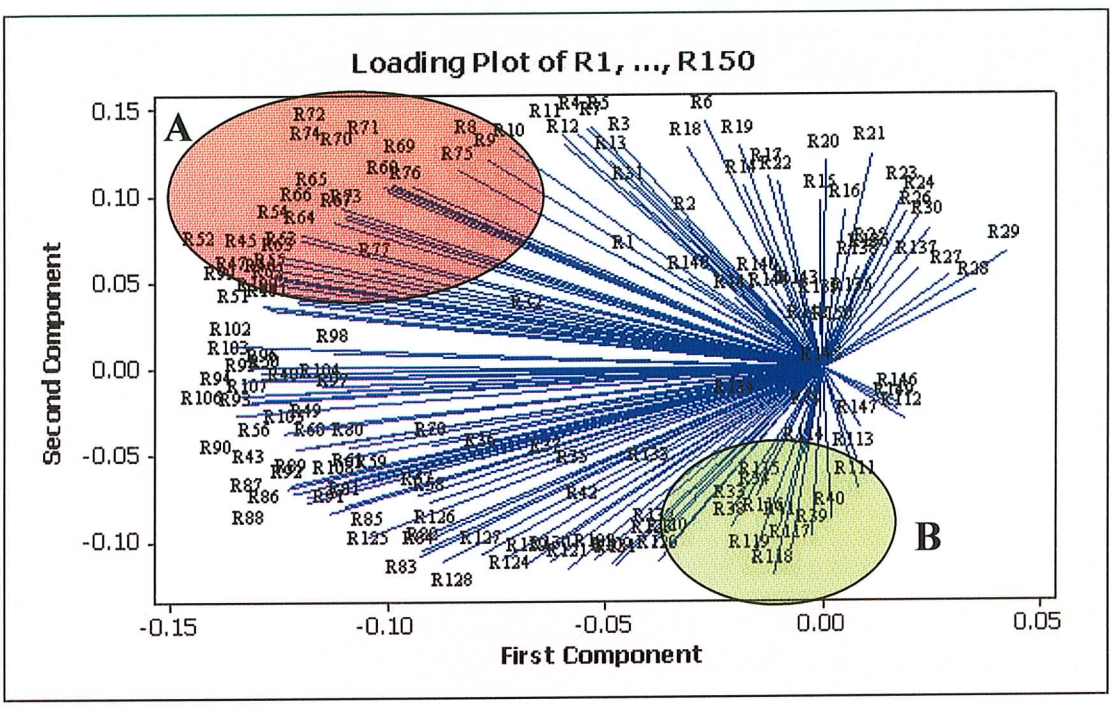


Figure 7.17 Loading plot for the PCs from 10 point average over entire spectral range

Other R-values strongly localised in the region associated with normal spectral features (Fig. 7.17A) are a large group ranging from R63 – R77 (1020 – 1170 cm^{-1}). It is clear from Figure 6.13 that this is the result of much stronger intensity values in this region in the spectra of normal tissue, compared to abnormal tissue. It is difficult to assign these bands to any specific Raman vibrational modes, due to the low resolution and the fact that many different bands overlap over this range. Other obvious features of interest include the Amide I band (R126 – R128 / 1650 – 1680 cm^{-1}) and symmetric ring breathing (R61 / 1000 – 1010 cm^{-1}), both of which are to be found equidistant from the two quadrants. These bands have therefore been shown to have similar characteristics in both normal and

abnormal tissues. This is in line with previous results, where no discernable differences could be seen in these bands.

Figure 7.18 shows the results of the discriminant analysis for the 10 point average over the entire spectral range and Table 7.5 gives the results for cross validation (488 correct assignments out of 498). Using the entire spectrum (even at a much reduced resolution) has resulted in a vast improvement in the classification sensitivity of the LDA model. 4 normal samples were misclassified as abnormal, 4 CIN samples were classified as invasive carcinoma and 2 invasive carcinoma samples were classified as CIN. Critically, no abnormal samples were classified as normal.

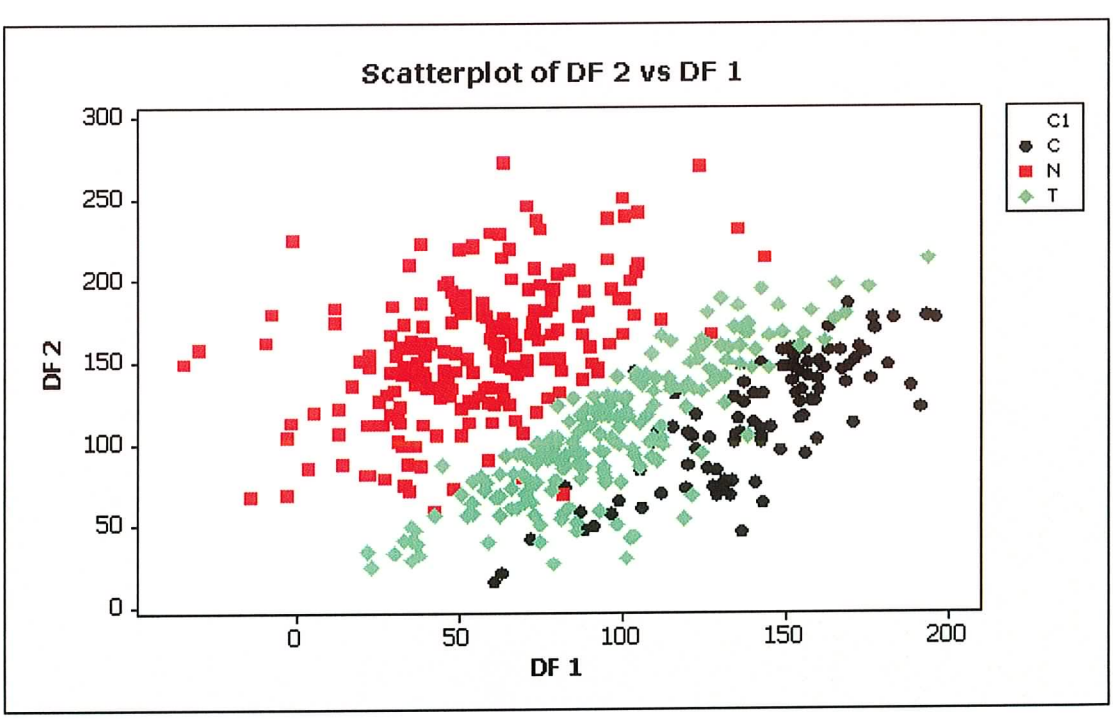


Figure 7.18 LDA plot generated from PCs for 10 point average over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	96	1	2
N	0	196	0
T	4	3	196
Total N	100	200	198
N correct	96	196	196
Proportion	0.960	0.980	0.990

N = 498
N Correct = 488
Proportion Correct = 0.980

Table 7.5 Results of classification for 10 point average over entire spectral range

As first derivative calculations were found to improve the previous results, the first derivative was calculated for the 10 point averaged spectra. This was again applied to the PCA analysis and the resulting scatterplot is shown in Figure 7.19.

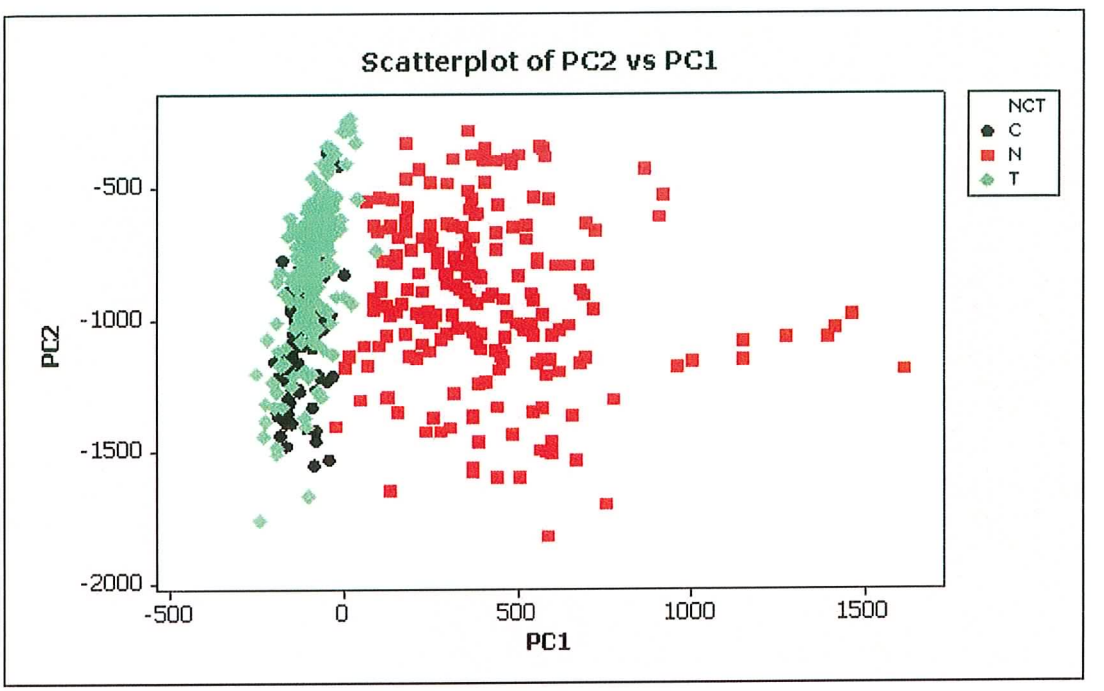


Figure 7.19 PCs scatterplot of 1st derivative of 10 point average over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

The LDA results are shown in Figure 7.20. It is clear that three distinct groups have been identified, corresponding to normal, CIN and invasive carcinoma. Table 7.6 gives the cross validation results and 488 of the 498 spectra have been again classified correctly. However, it is clear that there is a much better separation (Fig. 7.20) between the groups than was seen without derivation (Fig. 7.18).

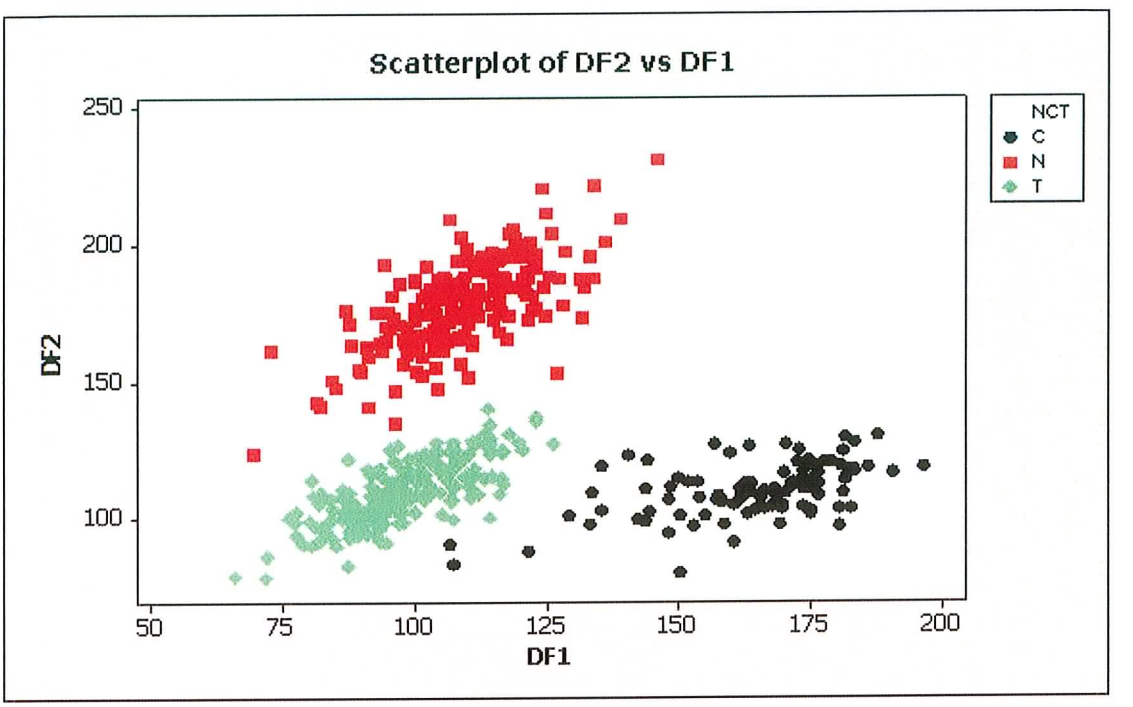


Figure 7.20 LDA of PCs of 1st derivative of 10 point average over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	94	0	1
N	0	197	0
T	6	3	197
Total N	100	200	198
N correct	94	197	197
Proportion	0.940	0.985	0.995

N = 498
N Correct = 488
Proportion Correct = 0.980

Table 7.6 Results of classification for 1st derivative of 10 point average over entire spectral range

7.4.4 Entire Range - 10 point average and Normalised

To examine whether normalising the data leads to an improvement, the 10 point averaged spectra were then normalised to the spectral maximum, from 0 to 1 and again subjected to PCA analysis (Fig. 7.21). The corresponding loading plot is shown in Figure 7.22.

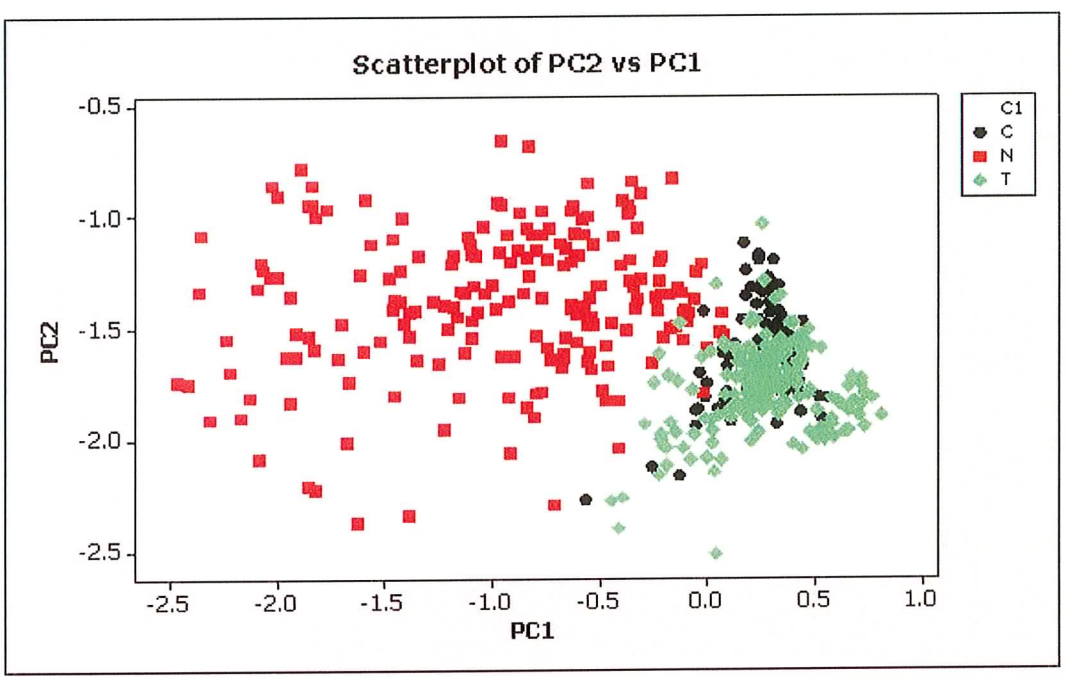


Figure 7.21 PCs scatterplot of normalised, 10 point averaged spectra, over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

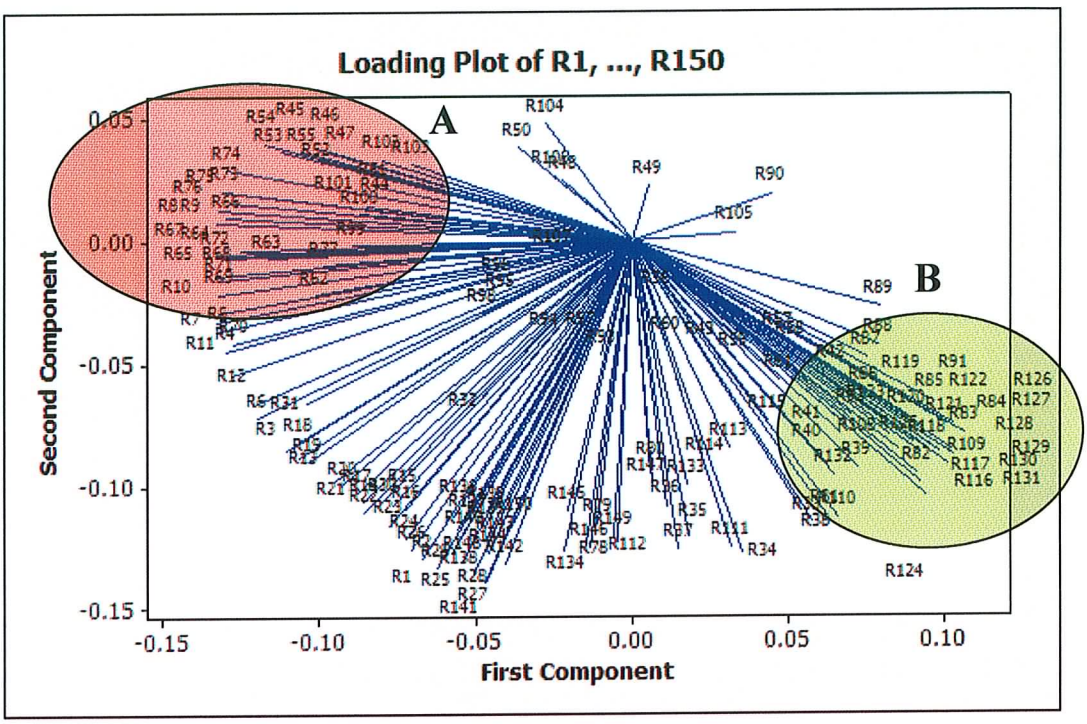


Figure 7.22 Loading plot for PCs of the 10 point averaged, normalised data

The loading plot of the normalised data differs in appearance from the loading plot of the un-normalised data (Fig. 7.17). Normalising redefines the principal components and has separated the groups more distinctively. This is apparent by separation along almost diagonally opposed quadrants.

PCs obtained were then applied to the LDA model (Fig. 7.23). Again all three tissue classes were identified and 492 out of 498 tissue spectra were correctly classified (Table 7.7). The cross validation mis-assigned 6 spectra, 2 of which were normal specimens assigned as invasive carcinoma. The other 4 were either invasive carcinoma or CIN mis-assigned as either CIN or invasive carcinoma respectively. Importantly no abnormal samples were classified as normal.

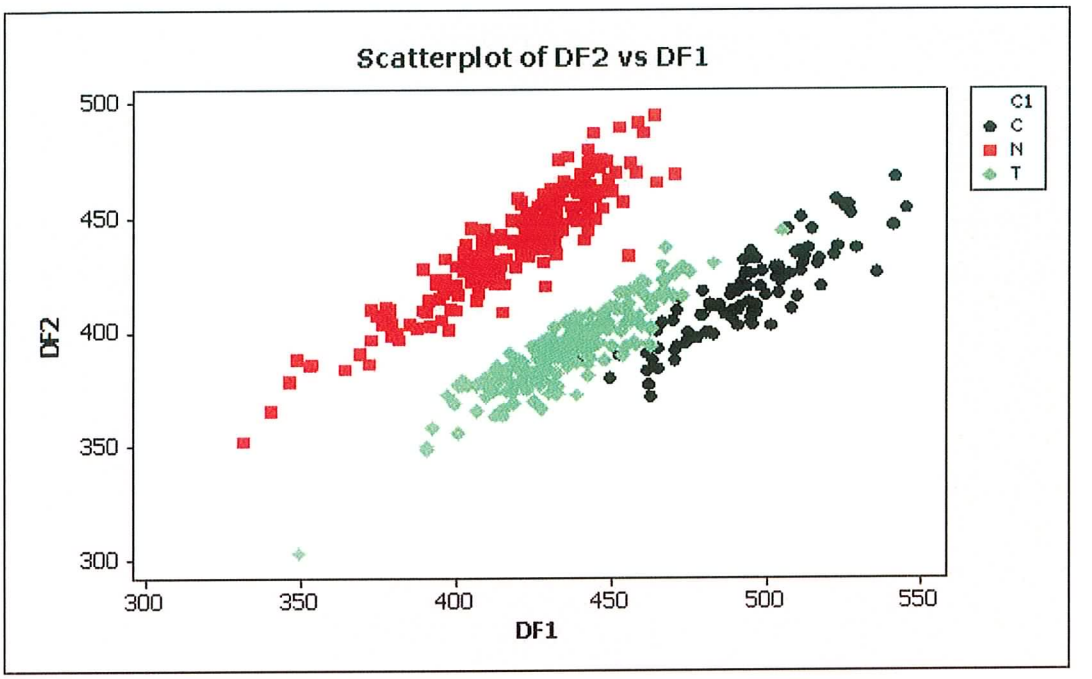


Figure 7.23 LDA of PCs of normalised, 10 point averaged spectra, over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	99	0	3
N	0	198	0
T	1	2	195
Total N	100	200	198
N correct	99	198	195
Proportion	0.990	0.990	0.985

N = 498
N Correct = 492
Proportion Correct = 0.988

Table 7.7 Results of classification for normalised, 10 point averaged spectra, over entire spectral range

The 1st derivative was also taken of the normalised, 10 point averaged, spectra. The scatterplot of the PCs is shown in Figure 7.24.

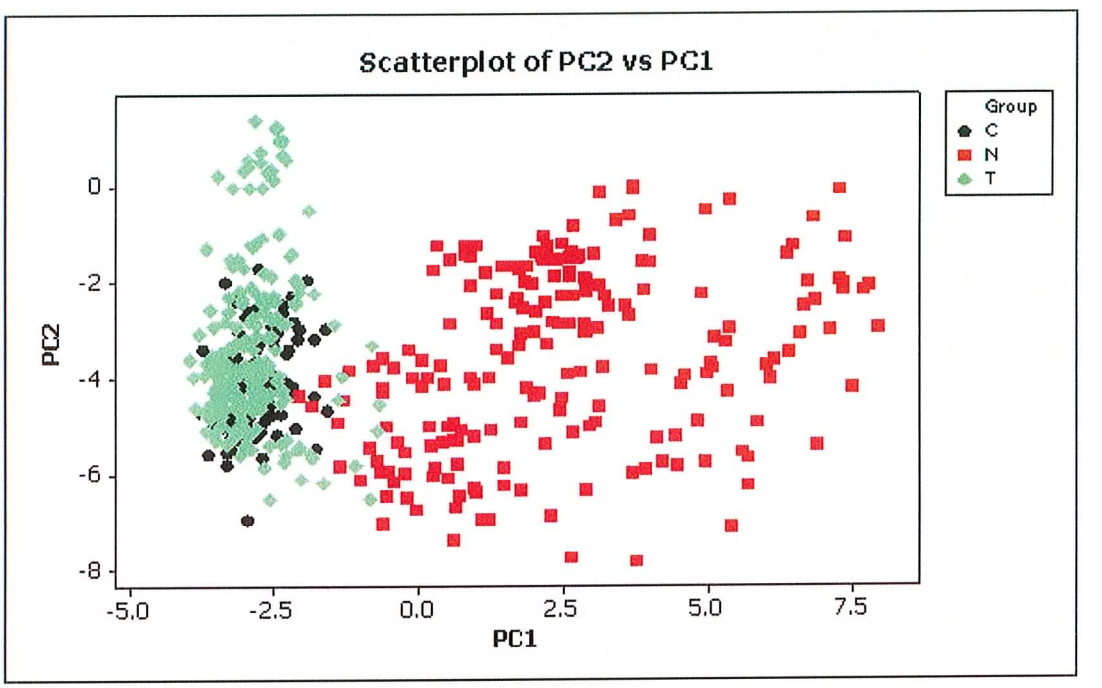


Figure 7.24 PCs scatterplot of 1st derivative, normalised, 10 point averaged spectra, over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

The LDA plot of these PCs is shown in Figure 7.25. The associated cross validation results are shown in Table 7.8. Again, 492 of the 498 spectra have been correctly classified. However, taking the 1st derivative of the spectra has resulted in a single abnormal specimen being labelled as normal.

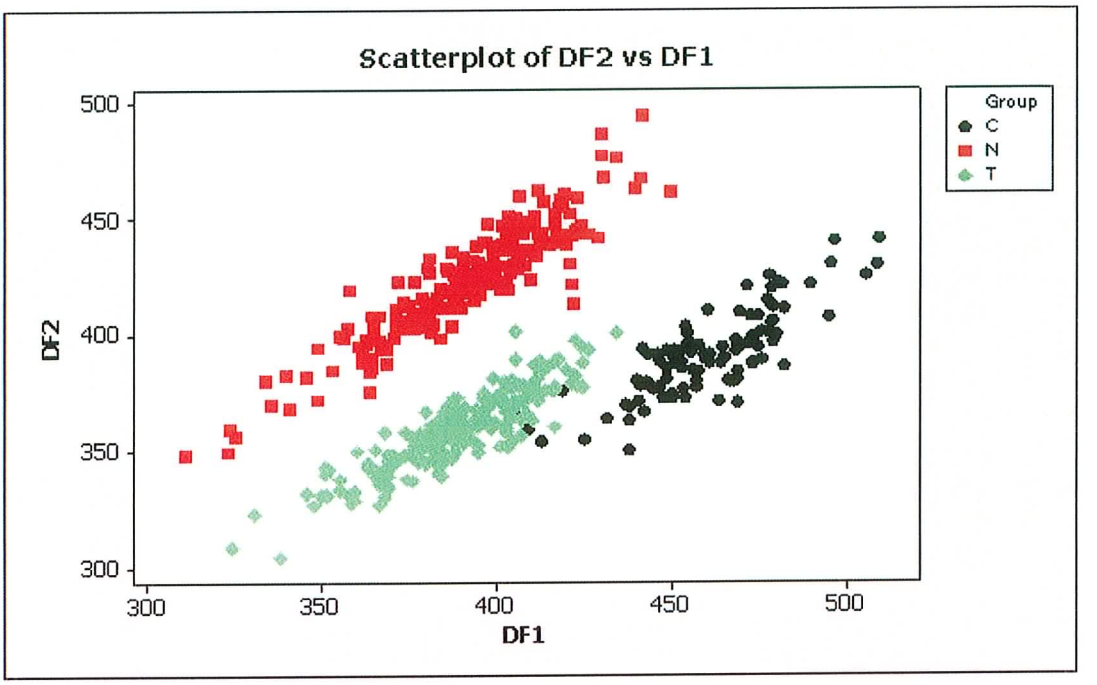


Figure 7.25 LDA of PCs of 1st derivative, normalised, 10 point averaged spectra, over entire spectral range (where, **C** = CIN, **N** = Normal and **T** = Invasive carcinoma)

Summary of Classification with Cross-validation

	True Group		
Put into Group	C	N	T
C	99	0	3
N	0	199	1
T	1	1	194
Total N	100	200	198
N correct	99	199	194
Proportion	0.990	0.995	0.980

N = 498
N Correct = 492
Proportion Correct = 0.988

Table 7.8 Results of classification for 1st derivative, normalised, 10 point averaged spectra, over entire spectral range

Values of sensitivity and specificity were then calculated (as outlined in section 7.2.6) for each of the different parameters examined (Table 7.9).

	NORMAL		CIN		INVASIVE CARCINOMA	
	SEN.	SPEC.	SEN.	SPEC.	SEN.	SPEC.
Specific Wavenumbers Normalised	92%	98.6%	75%	87.6%	82.3%	93.3%
Specific Wavenumbers Normalised - 1st derivative	92%	96.3%	90%	94.7%	92.2%	97.3%
Specific Wavenumbers Normalised - 2nd derivative	70.5%	85.8%	91%	92.2%	69.7%	82.3%
$10cm^{-1}$ Full Range	98%	100%	96%	99.2%	99%	97.6%
$10cm^{-1}$ Full Range 1st derivative	99.5%	100%	94%	99.7%	98.5%	97%
$10cm^{-1}$ Full Range Normalised	***99%***	***100%***	***99%***	***99.2%***	***98.5%***	***99%***
$10cm^{-1}$ Full Range Normalised 1st derivative	99.5%	99.6%	99%	99.2%	98%	99.3%

Table 7.9 Sensitivity (SEN.) and specificity (SPEC.) values achieved for different analysis parameters investigated

As shown in Table 7.9 the best values of sensitivity were obtained when the 10 cm^{-1} averaged spectra were applied to PCA and LDA. Using only narrow wavenumber regions produced sensitivity and specificity values of 92% and 98.6% for normal tissue, 75% and 87.6% for CIN and 82.3% and 93.3% for invasive carcinoma. Although these values are a credible result in themselves, there is clearly room for improvement. Taking the first derivative improved the sensitivity and specificity values for the CIN and invasive carcinoma tissues with values of 92% and 96.3% (normal), 90% and 94.7% (CIN) and 92.2% and 97.3% (invasive carcinoma) respectively. However taking the second derivative resulted in a decrease in sensitivity and specificity values, down to 70.5% and 85.8% (normal), 91% and 92.2% (CIN) and 69.7% and 82.3% (invasive carcinoma). By utilising the entire spectral range (with 10 cm^{-1} averaging) there is an immediate jump in the sensitivity and specificity values, with all tissue classes having values in the high 90's. The most successful classification results were achieved using the normalised full data range, with sensitivity and specificity values of 99% and 100% (normal tissue), 99% and 99.2% (CIN) and 98.5% and 99% (invasive carcinoma). The 100% specificity value achieved using this analysis method is particularly important, as it shows no abnormal samples being classified as normal i.e. no cancer or pre-cancerous samples would be missed during screening of these samples. Also as each patient in the study was classified by 10 spectra (as opposed to a single spectrum) the probability of incorrect classification is reduced further. The derivative of the 10 cm^{-1} averaged normalised spectra did not lead to further improvements in classification and is likely due to the fact that the derivative of a Gaussian is the same regardless of height, hence derivatisation performs the same function as normalisation. All different data analysis methods resulted in a greater differentiation between normal and abnormal tissue, than between CIN and invasive carcinoma tissue.

This is to be expected given the obvious similarity seen between the spectra of CIN and invasive carcinoma tissue as shown in the previous Chapter. It is the difference between normal and abnormal tissue that is the most important to detect.

7.5 Conclusion

PCA has proved to be an effective method of data reduction and analysis. This method alone produced good separation between tissue classes. When LDA was applied to the PCs excellent group separation was achieved. The benefit of using the entire spectrum was clear by the large range of markers which contribute to the PCs. The most successful classification results were achieved using the normalised full data range, with sensitivity and specificity values of 99% and 100% (normal tissue), 99% and 99.2% (CIN) and 98.5% and 99% (invasive carcinoma). Markers associated with glycogen, as well as broad features at 1020 – 1170 cm^{-1}, were found to separate normal tissue from abnormal tissue. Both CIN and invasive carcinoma tissue were found to have strong nucleic acid contributions. The first derivative of the spectra was found to improve group separation, although normalising the spectra was also found to have a similar effect. Taking the second derivative was found to produce a deterioration in classification efficiency, likely due to over sampling of the spectra. Due to limited availability of samples only one spectrum was omitted in the leave-one-out cross validation, if more samples were available it would have been preferable to leave one entire sample out when performing cross validation. Also validation on an entirely independent sample set in the future would be desirable. Sensitivity and specificity values compare well and often exceeded previously published results particularly with respect to cervical cancer (Mahadaven-Jansen *et al.* 1998).

Chapter 8

General Conclusion

The aim of this study was to evaluate the potential of vibrational spectroscopy for the diagnosis of human tumours. Both Raman and FTIR spectroscopies were employed. Initial investigations examined the components of human tissue. Spectra were recorded from amino acids, purines, pyrimidines, peptides, biological macromolecules (such as glycogen, collagen, DNA and RNA) and proteins. These spectra were used as a reference library to gain a better understanding of the complex features associated with the spectra of tissue. A variety of different tissues were examined, although cervical tissue was chosen for use in the cancer study. This was primarily due to the availability of pathologically characterised tissue through collaboration with the National Maternity Hospital, Dublin.

In order to validate the use of dewaxed FFPP tissue sections in the study, the effects of tissue processing on these sections were investigated. Both Raman and FTIR spectra were recorded from each of the processing steps the tissue underwent. It was found that Raman spectroscopy produced more detailed spectra and hence was more sensitive to changes in tissue composition that FTIR spectroscopy. Formalin fixation was found to produce changes in protein conformation, as well as a new Raman peak at 1490cm^{-1} likely due to the effects of methylene bridge formation. The study highlighted problems associated with residual wax after clearing. FFPP tissue sections, although reduced in overall intensity compared to fresh tissue, were found to retain sufficient biochemical similarity to fresh tissue to be diagnostically useful.

In order to achieve an improvement in the level of wax removal a study was carried out into a number of different commonly used dewaxing techniques. The study found that standard dewaxing protocols result in quantities of residual wax being present in the tissue after

dewaxing. Considerable improvements were made to the level of wax removal with the use of hexane as a dewaxing solvent. The use of hexane was also found to produce improved immunohistochemical staining. The present study recommends the use of hexane as a dewaxing agent.

FTIR spectroscopy was found to produce excellent separation between normal and invasive carcinoma tissue spectra. Spectra produced compared very well to previously published results. Different cell types were found to have their own distinct IR spectra, allowing cell identification based on their spectral signature. Normal, invasive carcinoma and CIN samples were examined. Dramatic reduction in glycogen band intensity (1152, 1079 and 1036cm^{-1}), as well as an increase in the nucleic acid band intensity (1236cm^{-1}) with the onset of carcinogenesis were observed.

In comparison to FTIR, Raman spectroscopy gives relatively weak spectra with longer accumulation times required. However, Raman spectroscopy has much higher spatial and spectral resolution, allowing better definition of biochemical composition and differentiation between cellular components. It has the added advantage of minimal interference from water.

The Raman spectra recorded showed specific spectral characteristics of different cells found in the cervix. Ectocervical epithelial cells, basal cells, connective tissue and endocervical columnar cells all display distinct spectral properties. Clear differences were observed between normal epithelium and invasive carcinoma. The strong glycogen bands associated with normal epithelium (482, 849, 938, 1082 and 1336 cm^{-1}) were absent or

significantly reduced in the spectra of invasive carcinoma. Invasive carcinoma and CIN tissues displayed prominent increases in intensity of bands at 724, 779 and 1578 cm^{-1}, but also increases in bands at 829, 852, 1002, 1098, 1240 cm^{-1}, mainly attributable to increases in nucleic acid concentrations.

To quantify the prediction accuracy associated with the Raman spectral data, multivariate methods were applied. A number of different "pre-treatment" steps were investigated including, highlighting specific peaks, normalising the data, taking a 10 cm^{-1} average of the spectra and taking the 1^{st} and 2^{nd} derivative spectra. Highlighting 9 wavenumber regions alone was found to produce the poorest classification results, with sensitivities and specificities of 92% and 98.6% respectively for normal tissue, 75% and 87.6% respectively for CIN and 82.3% and 93.3% respectively for invasive carcinoma. The best classification accuracy was achieved when a 10 cm^{-1} average of spectra was taken and spectra were normalised. This produced sensitivity and specificity values of 99.5% and 100% respectively for normal tissue, 99% and 99.2% respectively for CIN and 98.5% and 99% respectively for invasive carcinoma. Of 298 invasive carcinoma and neoplasia spectra applied to the discriminant analysis, not a single spectrum was misclassified as normal.

Vibrational spectroscopy has clearly been shown to be a powerful and versatile technique, with great potential to produce real-time, clinically relevant data. This study has resulted in very clear and consistent spectral markers of cancer and neoplasia. Multivariate analysis is a powerful technique to discriminate between the many subtle spectral features of the spectra recorded. The importance of using the entire spectral range recorded was also shown. Sensitivity and specificity values achieved are far in excess of those achieved using

current clinical methods. The technique has potential as an analytic tool to further investigate the biochemical mechanisms associated with a disease, in addition to its obvious diagnostic capabilities. With respect to cervical cancer, the challenge now is to examine different grades of CIN with a view to correct classification within this class and to identify the earliest possible biochemical markers of disease. The potential of the technique is not limited to cervical cancer, or even cancer. The future of vibrational spectroscopy as a diagnostic technique is potentially unlimited.

References

American Accreditation HealthCare Commission (2004) http://www.nlm.nih.gov/ medlineplus/encyclopedia.html

Arrondo JLR, Muga A, Castresana J, Goñi FM (1993) "Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy" Prog. Biophys. Molec. Biol. 59, 23-56

Bakker Schut TC, Wolthius R, Caspers PJ, Puppels GJ (2002) " Real-time tissue characterization on the basis of *in vivo* Raman spectra" Journal of Raman Spectroscopy, 33 580-585

Bancroft JD, Gamble M (2002) Theory and Practice of Histological Techniques, 5th Edition, Churchill Livingstone (Edinburgh)

Bankfalvi A, Navabi H, Bier B, Bocker W, Jasani B, Schmid KW (1994) "Wet autoclave pretreatment for antigen retrieval in diagnostic immunohistochemistry" Journal of Pathology 174 (3): 223-228

Barry BW, Edwards HGM, Williams AC (1992). "Fourier-Transform Raman and Infrared Vibrational Study of Human Skin -Assignment of Spectral Bands" Journal of Raman Spectroscopy 23(11): 641-645.

Bindig U, Frank F, Gersonde I, Meinke M, Zelianeos K, Katzir A, Muller G (2003) "Fiber-optic and microscopic infrared biodiagnostics" Laser Physics 13 (1): 96-105

Boere IA, Schut TCB, van den Boogert J, de Bruin RWF, Puppels GJ (2003) "Use of fibre optic probes for detection of Barrett's epithelium in the rat oesophagus by Raman spectroscopy" Vibrational Spectroscopy 32 (1): 47-55

Bonin S, Petrera F, Niccolini B, Stanta G (2003) "PCR analysis in archival postmortem tissues" Journal of Clinical Pathology-Molecular Pathology 56 (3): 184-186

Bozec L, Hammiche A, Tobin MJ, Chalmers JM, Everall NJ, Pollock HM (2002) "Near-field photothermal Fourier transform infrared spectroscopy using synchrotron radiation" Measurement Science & Technology 13 (8): 1217-1222

Brereton R.G. (2005) Chemometrics – Data Analysis for the Laboratory and Chemical Plant, John Wiley & Sons, Chichester, England

Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, Feld MS (2001) "Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ" Cardiovascular Pathology 10 (2): 69-82

Buschman HP, Marple ET, Wach ML, Bennett B, Schut TCB, Bruining HA, Bruschke AV, Van der Laarse A, Puppels GJ (2000) "In vivo determination of the molecular composition

of artery wall by intravascular Raman spectroscopy" Analytical Chemistry 72 (16): 3771-3775

Caspers PJ, Lucassen GW, Carter EA, Bruining HA, Puppels GJ (2001). "In Vivo Confocal Raman Microspectroscopy of the Skin: Noninvasive Determination of Molecular Concentration Profiles." The Journal of Investigative Dermatology 116(3): 434-442.

Caron M. (2004)http://apps.bioneq.qc.ca/twiki/bin/view/Basedeconnaissances/FormatsDeFichiers

Cattoretti G, Pileri S, Parravicini C, Becker MHG, Poggi S, Bifulco C, Key G, Damato L, Sabattini E, Feudale E, Reynolds F, Gerdes J, Rilke (1993) F "Antigen unmasking on formalin-fixed, paraffin-embedded tissue-sections" Journal of Pathology 171 (2): 83-98

Chan PKS, Chan DPC, To KF, Yu MY, Cheung JLK, Cheng AF (2001) "Evaluation of extraction methods from paraffin wax embedded tissues for PCR amplification of human and viral DNA" Journal of Clinical Pathology 54 (5): 401-403

Chang JI, Huang YB, Wu PC, Chen CC, Huang SC, Tsai YH (2003) "Characterization of human cervical precancerous tissue through the Fourier transform infrared microscopy with mapping method" Gynecologic Oncology 91 (3): 577-583

Chiriboga L, Xie P, Yee H, Vigorita V, Zarou D, Zakim D, Diem M (1998 a) "Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelial cells in the human cervix" Biospectroscopy 4(1):47-53

Chiriboga L, Xie P, Vigorita V, Zarou D, Zakim D, Diem M (1998 b) "Infrared spectroscopy of human tissue. II. A comparative study of spectra of biopsies of cervical squamous epithelium and of exfoliated cervical cells" Biospectroscopy. 1998; 4 (1): 55-9

Chiriboga L, Xie P, Zhang W, Diem M (1998 c) "Infrared spectroscopy of human tissue. III. Spectral differences between squamous and columnar tissue and cells from the human cervix" Biospectroscopy 3, 253-256

Chiriboga L, Yee M, Diem M (2000) "Infrared spectroscopy of human cells and tissue. Part VI: A comparative study of histopathology and infrared microspectroscopy of normal, cirrhotic and cancerous liver tissue" Applied Spectroscopy 54 (1): 1-8

Choo-Smith LP, Maquelin K, van Vreeswijk T, Bruining HA, Puppels GJ, Thi NAG, Kirschner C, Naumann D, Ami D, Villa AM, Orsini F, Doglia SM, Lamfarraj H, Sockalingum GD, Manfait M, Allouch P, Endtz HP (2001) "Investigating microbial (micro)colony heterogeneity by vibrational spectroscopy" Applied and Environmental Microbiology 67 (4): 1461-1469

Cohenford MA, Rigas B (1998) "Cytologically normal cells from neoplastic cervical samples display extensive structural abnormalities on IR spectroscopy: implications for tumor biology" Proc Natl Acad Sci U S A;95(26):15327-32.

Cohenford MA, Godwin TA, Cahn F, Bhandare P, Caputo TA, Rigas B. (1997) "Infrared spectroscopy of normal and abnormal cervical smears: evaluation by principal component analysis." Gynecol Oncol. 66(1):59-65

Crow P, Stone N, Kendall C, Uff J, Ritchie A, Wright M (2003) "Optical diagnostics in bladder cancer: The use of Raman spectroscopy to differentiate between bladder pathologies in vitro" Journal Of Urology 169 (4): 871 Suppl. S

Crow P, Kendall C, Uff J, Gilbert H, Persad R, Wright M (2003) "Optical diagnostics in prostate cancer: The use of Raman spectroscopy to differentiate between different prostatic pathologies in vitro" Journal of Urology 169 (4): 1618 Suppl. S

Culling CFA, Allison RT, Barr WT (1985) Cellular Pathology Technique, 4th edition, Butterworth's

De Jong BWD, Schut TCB, Coppens J, Wolffenbuttel KP, Kok DJ, Puppels GJ (2003) "Raman spectroscopic detection of changes in molecular composition of bladder muscle tissue caused by outlet obstruction" Vibrational Spectroscopy 32 57-65

De Jong BWD, Schut TCB, Wolffenbuttel KP, Nijman JM, Kok DJ, Puppels GJ (2002) "Identification of bladder wall layers by Raman spectroscopy" Journal of Urology 168 (4): 1771-1778 Part 2

Diem M, Romeo M, Boydston-White S, Miljkovic M, Matthaus C (2004) "A decade of vibrational micro-spectroscopy of human cells and tissue (1994-2004)" Analyst 129 (10): 880-885

Diem M, Chiriboga L, Lasch P, Pacifico A (2002) "IR spectra and IR spectral maps of individual normal and cancerous cells" Biopolymers 67 (4-5): 349-353

Dovbeshko GI, Gridina NY, Kruglova EB, Pashchuk OP (2000) "FTIR spectroscopy studies of nucleic acid damage" Talanta 53 (1): 233-246 Sp. Iss. SI

Drury RAB, Wallington EA (1980) "Carleton's Histological Technique" Fifth Edition, Oxford University Press.

Duffy B (2002) Development of Novel Gas Sensors, DCU, Dublin, PhD Thesis

Dukor R (2002) Handbook of Vibrational Spectroscopy, Volume 5

Eriksson L, Johansson E, Kettaneh-Wold N, Wold S (2001) "Multi- and Megavariate Data Analysis –principals and applications" Umtrics Academy, Volume 1

Fung MFK, Senterman M, Eid P, Faught W, Mikhael NZ, Wong PTT (1997) "Comparison of Fourier-Transform Infrared Spectroscopic Screening of Exfoliated Cervical Cells with Standard Papanicolaou Screening" Gynecol. Oncol., 66, 10

Fung MFK, Senterman MK, Mikhael NZ, Lacelle S, Wong PTT (1996) "Pressure-tuning Fourier transform infrared spectroscopic study of carcinogenesis in human endometrium" Biospectroscopy 2 (3): 155-165

Gniadecka M, Wulf HC, Nielsen OF, Christensen DH, Hercogova J (1997) "Distinctive molecular abnormalities in benign and malignant skin lesions: Studies by Raman spectroscopy" Photochemistry and Photobiology 66 (4): 418-423

Gowri V, (2000) "Uterine Inversion and Corpus Malignancies: A Historical Review" Obstetrical & Gynecological Survey. 55(11):703-707

Harcourt Grace and Company (1999) http://www.people.virginia.edu/~rjh9u/prot2ndstruct .html

Hawi SR, Cambell WB, Kajdacsy-Balla A, Murphy R, Adar F, Nithipatikom K (1996). "Characterization of normal and malignant hepatocytes by Raman Microspectroscopy." Cancer Letters 110: 35-40

Huang ZW, McWilliams A, Lui H, McLean DI, Lam S, Zeng HS (2003) "Near-infrared Raman spectroscopy for optical diagnosis of lung cancer" International Journal of Cancer 107 (6): 1047-1052

Kaiser G.E. (2005) http://student.ccbcmd.edu/~gkaiser/goshp.html

Kaminaka S. ; Yamazaki, H. I., T.; Kohda, E.; Hamaguchi, H. (2001). "Near-Infrared Raman Spectroscopy of human lung tissues; possibility of molecular-level cancer diagnosis." Journal of Raman Spectroscopy 32: 139-141

Keller S, Schrader B, Hoffmann A, Schrader W, Metz K, Rehlaender A, Pahnke J, Ruwe M, Budach W (1994) "Application of Near-Infrared Fourier-Transform Raman-Spectroscopy in Medical-Research" Journal Of Raman Spectroscopy 25 (7-8): 663-671

Kendall C, Stone N, Shepherd N, Geboes K, Warren B, Bennett R, Barr H (2003) "Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus" Journal of Pathology 200 (5): 602-609

Kneipp J, Schut TB, Kliffen M, Menke-Pluijmers M, Puppels G (2003) "Characterization of breast duct epithelia: a Raman spectroscopic study" Vibrational Spectroscopy 32 (1): 67-74

Kneipp J, Miller LM, Joncic M, Kittel M, Lasch P, Beekes M, Naumann D (2003b) "In situ identification of protein structural changes in prion-infected tissue" Biochimica Et Biophysica Acta-Molecular Basis of Disease 1639 (3): 152-158

Knudsen L, Johansson CK, Philipsen PA, Gniadecka M, Wulf HC (2002) "Natural variations and reproducibility of in vivo near-infrared Fourier transform Raman spectroscopy of normal human skin" Journal of Raman Spectroscopy 33 (7): 574-579

Koljenovic S, Choo-Smith LP, Schut TCB, Kros JM, van den Berge HJ, Puppels GJ (2002) "Discriminating vital tumor from necrotic tissue in human glioblastoma tissue samples by Raman spectroscopy" Laboratory Investigation 82 (10): 1265-1277

Krzanowski M.J., (1988) Principals of Multivariate Analysis, Oxford Statistical Science series, Clarendon Press, New York,

Lakshmi RJ, Kartha VB, Krishna CM, Solomon JGR, Ullas G, Devi PU, (2002)" Tissue Raman spectroscopy for the study of radiation damage: Brain irradiation of mice" Radiation Research 157(2): 175–182

Lau DP, Huang Z, Lui H anderson DW, Berean K, Morrison MD, Liang S, Zeng H "Raman spectroscopy for optical diagnosis in the larynx: preliminary findings" (2005) Lasers in Surgery and Medicine, 37 192-200

Lau DP, Huang Z, Lui H, Man CS, Berean K, Morrison MD, Zeng H (2003) "Raman spectroscopy for optical diagnosis in normal and cancerous tissue of the nasopharynx – preliminary findings" Lasers in Surgery and Medicine, 32 210-214

Levin IW, Bhargava R (2005) "Fourier transform infrared vibrational spectroscopic imaging: Integrating microscopy and molecular recognition" Annual Review of Physical Chemistry 56: 429-474

Li QB, Xu Z, Xu YZ, Zhang YF, Zhang NW, Wang LX, Sun XJ, Zhang L, Wang F, Yang LM, Zhao Y, Ren Y, Liu Z, Weng SF, Zhou WJ, Wu JG (2004) "Application of Fourier transform infrared spectroscopy to non-invasive detection of breast tumor in vivo and in situ" Chemical Journal of Chinese Universities-Chinese 25 (11): 2010-2012

Ling XF, Xu Z, Xu YZ, Li QB, Zhou S, Zhang L, Zhao HM, Hou CS, Wang LX, Hou KY, Zhou XS, Wu JG (2005) "FTIR spectroscopic explorations of clinical practice of breast cancer" Spectroscopy and Spectral Analysis 25 (2): 198-200

Liu G, Liu JH, Zhang L, Yu F, Sun SZ (2005), "Raman spectroscopic study of human tissues" Spectroscopy and Spectral Analysis, 25 (5) 723-725

Liu KZ, Schultz CP, Salamon EA, Man A, Mantsch HH (2003) "Infrared spectroscopic diagnosis of thyroid tumors" Journal of Molecular Structure 661: 397-404

Liu CH, Das BB, Glassman WLS, Tang GC, Yoo KM, Zhu HR, Akins DL, Lubicz SS, Cleary J, Prudente R, Celmer E, Caron A, Alfano RR (1992) "Raman, Fluorescence and Time-Resolved Light-Scattering As Optical Diagnostic-Techniques To Separate Diseased And Normal Biomedical Media" Journal of Photochemistry and Photobiology B-Biology 16 (2): 187-209

Liu KZ, Jackson M, Sowa MG, Ju HS, Dixon IMC, Mantsch HH (1996) "Modification of the extracellular matrix following myocardial infarction monitored by FTIR spectroscopy" Biochimica et Biophysica Acta-Molecular Basis Of Disease 1315 (2): 73-77

London Science Museum (2005) http://www.sciencemuseum.org.uk/exhibitions/lifecycle /74.asp

MacIntyre N (2001) "Unmasking antigens for immunohistochemistry" British Journal of Biomedical Science 58 (3): 190-196

Mahadevan-Jansen A., Ramanujam N., Malpica A., Thomsen S., Utzinger U., Richards-Kortum R. (1998). "Near-Infrared Raman Spectroscopy for In Vitro Detection of Cervical Precancers." Photochemistry and Photobiology 68(1): 123-132.

Mallery C (2005) http://fig.cox.miami.edu/~cmallery/150/protein/proteinsb.htm

Manoharan R, Wang Y, Feld MS (1996) "Histochemical analysis of biological tissues using Raman spectroscopy" Spectrochimica Acta Part A-Molecular and Biomolecular Spectroscopy 52 (2): 215-249

Mantsch HH, Choo-Smith LP, Shaw RA (2002) "Vibrational spectroscopy and medicine: an alliance in the making" Vibrational Spectroscopy 30 (1): 31-41

McIntosh LM, Jackson M, Mantsch HH, Mansfield JR, Crowson AN, Toole JWP (2002) "Near-infrared spectroscopy for dermatological applications" Vibrational Spectroscopy 28 (1): 53-58

Miller LM, Carr GL, Williams GP, Chance MR (1997) "Synchrotron infrared microspectroscopy as a means of studying chemical composition at a cellular level" Biophysical Journal 72 (2): TU445-TU445 Part 2

Mizuno A, Hitajima K., Kawauchi K, Muraishi S, Ozaki Y (1994). "Near-Infrared Fourier-Transform Raman-Spectroscopic Study of Human Brain-Tissues and Tumors." Journal Of Raman Spectroscopy 25(1): 25-29.

Mohlenhoff B, Romeo M, Diem M, Wood BR (2005) "Mie-type scattering and non-Beer-Lambert absorption behavior of human cells in infrared microspectroscopy" Biophysical Journal 88 (5): 3635-3640

Molckovsky A, Song LMWK, Shim MG, Marcon NE, Wilson BC (2003) "Diagnostic potential of near-infrared Raman spectroscopy in the colon: differentiating adenomatous from hyperplastic polyps" Gastrointestinal Endoscopy 57 (3): 396-402

Mordechai S, Sahu RK, Hammody Z, Mark S, Kantarovich K, Guterman H, Podshyvalov A, Goldstein J, Argov S (2004) "Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma" Journal of Microscopy-Oxford 215: 86-91

Morris BJ, Lee C, Nightingale BN, Molodysky E, Morris LJ, Appio R, Sternhell S, Cardona M, Mackerras D, Irwig LM (1995) "Fourier-transform infrared-spectroscopy of dysplastic, papillomavirus-positive cervicovaginal lavage specimens" Gynecologic Oncology 56 (2): 245-249

National Cancer Registry (2005) "Cancer in Ireland (1994 – 2001)" Irish National Cancer Report

Naumann D (2001) "FT-infrared and FT-Raman spectroscopy in biomedical research" Applied Spectroscopy Reviews 36 (2-3): 239-298

Nedzel GA (1951) "Intranuclear birefringent inclusions, an artifact occurring in paraffin sections" Quart. Journal of Micr. Science 92, 343

Neviliappan S, Fang Kan L, Tiang Lee Walter T, Arulkumaran S, Wong PT. (2002) "Infrared spectral features of exfoliated cervical cells, cervical adenocarcinoma tissue and an adenocarcinoma cell line (SiSo)" Gynecol Oncol. 85(1):170-174

Nijssen A, Bakker Schut TC, Heule F, Caspers PJ, Hayes DP, Neumann MHA, Puppels GJ (2002). "Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy" Journal of Investigative Dermatology 119(1): 64-69

Norton AJ, Jordan S, Yeomans P (1994) "Brief, high-temperature heat denaturation (pressure cooking) - a simple and effective method of antigen retrieval for routinely processed tissues" Journal of Pathology 173 (4): 371-379

Notingher I, Verrier S, Haque S, Polak JM, Hench LL (2003) "Spectroscopic study of human lung epithelial cells (A549) in culture: living cells versus dead cells" Biopolymers 72(4):230-40

Nunes LD, Martin AA, Silveira L, Zampieri M (2003) "FT-Raman spectroscopy study for skin cancer diagnosis" Spectroscopy-An International Journal 17 (2-3): 597-602

Ó Faoláin E, Hunter MB, Byrne JM, Kelehan P, Lambkin HA, Byrne HJ, Lyng FM (2005) "A study examining the effect of tissue processing on human tissue sections using vibrational spectroscopy" Journal of Vibrational Spectroscopy. 53: 121-131

Ó Faoláin E, Hunter MB, Byrne JM, Kelehan P, Lambkin HA, Byrne HJ, Lyng FM (2005) "Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents" J. Histochem. Cytochem. 53: 121-131

Paluszkiewicz C, Kwiatek WM (2001) "Analysis of human cancer prostate tissues using FTIR microspectroscopy and SRIXE techniques" Journal of Molecular Structure 565: 329-334 Sp. Iss. SI

Pleshko NL, Boskey AL, Mendelsohn (1992) "An FT-IR microscopic investigation of the effects of tissue preservation on bone" R Calcified Tissue International 51 (1): 72-77

Polak J.M., Van Noorden S. (1997) "Introduction to Immunocytochemistry" Second Edition, BIOS Scientific Publishers Limited

Ren Y, Xu YZ, Wang J, Zhang YF, Wang F, Shi JS, Wu JG (2003) "FTIR spectroscopic and statistical studies on the lung tissues" Spectroscopy and Spectral Analysis 23 (4): 681-684

Romeo M, Matthaus C, Miljkovic M, Diem M (2004) "Infrared microspectroscopy of individual human cervical cancer (HeLa) cells" Biopolymers 74 (1-2) 168-171

Romeo MJ, Quinn MA, Burden FR, McNaughton D (2002a) "Influence of benign cellular changes in diagnosis of cervical cancer using IR microspectroscopy" Biopolymers 67 (4-5) 362-366

Romeo MJ, Wood BR, McNaughton D (2002b) "Observing the cyclical changes in cervical epithelium using infrared microspectroscopy" Vibrational Spectroscopy 28 167-175

Schurter MJ, LeBrun DP, Harrison KJ (2002) "Improved technique for fluorescence in situ hybridisation analysis of isolated nuclei from archival, B5 or formalin fixed, paraffin wax embedded tissue" Journal of Clinical Pathology-Molecular Pathology 55 (2): 121-124

Schut TCB, Wolthuis R, Caspers PJ, Puppels GJ (2002) "Real-time tissue characterization on the basis of in vivo Raman spectra" Journal of Raman Spectroscopy 33 (7): 580-585

Severcan F, Kaptan N, Turan B (2003) "Fourier transform infrared spectroscopic studies of diabetic rat heart crude membranes" Spectroscopy - An International Journal 17 (2-3): 569-577

Shafer-Peltier KE, Haka AS, Fitzmaurice M, Crowe J, Myles J, Dasari RR, Feld MS (2002 a) "Raman microspectroscopic model of human breast tissue: implications for breast cancer diagnosis in vivo" Journal of Raman Spectroscopy 33 (7): 552-563

Shafer-Peltier KE, Haka AS, Motz JT, Fitzmaurice M, Dasari RR, Feld MS "Model-based biological Raman spectral imaging" (2002 b) Journal of Cellular Biochemistry 125-137 Suppl. 39

Shah N, Cerussi A, Eker C, Espinoza J, Butler J, Fishkin J, Hornung R, Tromberg B (2001) "Noninvasive functional optical spectroscopy of human breast tissue" Proc. Nat. Acad. Sci. 98 (8): 4420-4425

Shaw RA, Guijon FB, Paraskevas M, Ying SL, Mantsch HH (1999) "Infrared spectroscopy of exfoliated cervical cell specimens. Proceed with caution" Anal. Quantitat. Cytol. Histol., 21, 292

Shi SR, Key ME, Kalra KL (1991) "Antigen retrieval in formalin-fixed, paraffin embedded tissues - an enhancement method for immunohistochemical staining based on microwave-oven heating of tissue-sections" Journal of Histochemistry & Cytochemistry 39 (6): 741-748

Shim MG, Wilson BC (1996) "The effects of ex vivo handling procedures on the near-infrared Raman spectra of normal mammalian tissues" Photochemistry and Photobiology 63 (5): 662-671

Sinauer (1998) Sinauer Associates http://www.mie.utoronto.ca/labs/lcdlab/biopic/fig/3.21.jpg

Sindhuphak R, Issaravanich S, Udomprasertgul V, Srisookho P, Warakamin S, Sindhuphak S, Boonbundarlchai R, Dusitsin N. (2003) "A new approach for the detection of cervical cancer in Thai women." Gynecologic Oncology, 90 (1) 10-14

Smith J, Kendall C, Sammon A, Christie-Brown J, Stone N (2003) "Raman spectral mapping in the assessment of axillary lymph nodes in breast cancer" Technology in Cancer Research & Treatment 2 (4): 327-331 2003

Socrates G, (2004) "Infrared and Raman Characteristic Group Frequencies : Tables and Charts - 3 edition" John Wiley & Sons

Song LMWK, Gao TY, Molckovsky A, Burgart L, Smyrk T, Buttar N, Lutzke L, Wilson B, Wang K (2003) "Accuracy of near-infrared Raman spectroscopy for differentiating adenocarcinoma from high-grade dysplasia in Barrett's esophagus" Gastrointestinal Endoscopy 57 (5): AB91-AB91 Suppl.

Spectral Database of Organic Compounds (2006), National Institute of Advanced Industrial Science and Technology, Japan http://www.aist.go.jp/RIODB/SDBS/

Stone N, Kendall C, Shepard N, Crow P, Barr H (2002) "Near-infrared Raman spectroscopy for the classification of epithelial and pre-cancers and cancer" Journal of Raman Spectroscopy (33) 564-573

Stone N, Stavroulaki P, Kendall C, Birchall M, Barr H (2000). "Raman spectroscopy for early detection of laryngeal malignancy: Preliminary results" Laryngoscope 110(10): 1756-1763.

Stryer L. (2005) Stryer's Biochemistry, W.H. Freeman & Company, 5th Edition

Tfayli A, Piot O, Durlach A, Bernard P, Manfait M (2005) "Discriminating nevus and melanoma on paraffin-embedded skin biopsies using FTIR micro spectroscopy" Biochimica et Biophysica Acta-General Subjects 1724 (3): 262-269

Thermo Electron Corporation (2006) http://www.thermo.com/com/cda/resources/resources_detail/1,2166,13428,00.html

Tobin M, Rutten F, Chesters M, Chalmers J, Symonds I, Fisher S, Allibone R, Hitchcock A (2002) "Investigating the Potential for Infrared Microanalysis in Cancer Screening" European Clinical Laboratory 20-22

Utzinger U, Mahadevan-Jansen A, Hinzelman D, Follen M, Richards-Kortum R (2001) "Near Infrared Raman Spectroscopy for In Vivo Detection of Cervical Precancers" Applied Spectroscopy 55 (8) 955

Vlachos JD (1968) "Birefringence and paraffinophilia of cell nuclei", Stain Technology, 43 (2) 89

Wang Q, Sanad W, Miller LM, Voigt A, Klingel K, Kandolf R, Stangl K, Baumann G (2005) "Infrared imaging of compositional changes in inflammatory cardiomyopathy" Vibrational Spectroscopy 38 (1-2): 217-222

Watson PF and Morris GJ (1987) "Cold shock injury in animal cells. In Temperature and Animal cells (K. Bowler, B.J. Fuller eds)" Symposia of the Society for Experimental Biology, 41, Company of Biologists, Cambridge, UK. 311-340

Wilbur DC (1997) "False negatives in focused rescreening of Papanicolaou smears: how frequently are 'abnormal' cells detected in retrospective review of smears preceding cancer or high-grade intraepithelial neoplasia" Archives of Pathology and Laboratory Medicine 121 (3) 273-6

Wojciechowski PM, Zierkiewicz W, Michalska D, Hobza P (2003) "Electronic structures, vibrational spectra and revised assignment of aniline and its radical cation: Theoretical study" Journal of Chemical Physics 118 (24): 10900-10911

Wold S, Sjöström M (1998) "Chemometrics, present and future success" Chemometrics and Intelligent Laboratory Systems (44) 3-14

Wong PTT, Senterman MK, Jackli P, Wong RK, Salib S, Campbell CE, Feigel R, Faught W, Fung MFK (2002) "Detailed account of confounding factors in interpretation of FTIR spectra of exfoliated cervical cells" Biopolymers 67 (6) 376-386

Wong PTT, Lacelle S, Fung MFK, Senterman M, Mikhael NZ (1995) "Characterization of exfoliated cells and tissues from human endocervix and ectocervix by FTIR and ATR/FTIR spectroscopy" Biospectroscopy 1 (5) 357-364

Wong PTT, Wong RK, Fung MFK (1993) "Pressure-Tuning FT-IR Study Of Human Cervical Tissues" Applied Spectroscopy 47 (7) 1058-1063

Wong PTT, Wong RK, Caputo TA, Godwin TA, Rigas B. (1991) "Infrared spectroscopy of exfoliated human cervical cells: evidence of extensive structural changes during carcinogenesis." Proceedings of the National Academy of Sciences 88(24) 10988-92

Wood BR, Chiriboga L, Yee H, Quinn MA, McNaughton D, Diem M (2003) "Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone and dysplastic squamous epithelium" Gynecologic Oncology 93 (1) 59-68

Wood BR, Quinn M, Tait B, Romeo M, McNaughton D (1998) "An FTIR Spectroscopic study to identify potential confounding variables and cell types in screening for cervical malignancies" Biospectroscopy (4) 75-91

Wood BR, Quinn MA, Burden FR, McNaughton D (1996) "An investigation into FTIR spectroscopy as a biodiagnostic tool for cervical cancer" Biospectroscopy (2) 143–153

Yazdi Y, Ramanujam N, Lotan R, Mitchell MF, Hittelman W, Richards-Kortum R (1998) "Resonance Raman spectroscopy at 257 nm excitation of normal and malignant cultured breast and cervical cells." Applied Spectroscopy 53 (1) 82-85

Yazdi HM, Bertrand MA, Wong PTT (1996) "Detecting structural changes at the molecular level with FT-IR spectroscopy" Acta Cytologica, (40) 664-668

Appendix I

Publications

Ó Faoláin E, Hunter MB, Byrne JM, Kelehan P, Byrne HJ, Lyng FM (2005) "The potential of vibrational spectroscopy in the early detection of cervical cancer: an exciting emerging field" Proc SPIE Vol. 5826 25-36

Ó Faoláin E, Hunter MB, Byrne JM, Kelehan P, McNamara M, Byrne HJ, Lyng FM (2005) "A study examining the effect of tissue processing on human tissue sections using vibrational spectroscopy" Journal of Vibrational Spectroscopy, 38/1-2 pp 121-127

Ó Faoláin E, Hunter MB, Byrne JM, Kelehan P, Lambkin HA, Byrne HJ, Lyng FM (2005) "Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents" J. Histochem. Cytochem. 53: 121-131

Ó Faoláin E., Conroy J., Hunter M., Kelehan P., Lambkin H., Mothersill C., Byrne H. and Lyng F.M (2002) "The potential of Raman spectroscopy for the molecular characterisation of human cervical tumours" SPIE Proceedings 4876 pp 84-91

Conroy J, Ó Faoláin E, Lyng F.M., Byrne H.J. (2002) "Raman spectroscopic analysis of ionization processes in biological systems" SPIE Proceedings 4876 pp 67-72

Selected Presentations

- Oral presentation 26th Annual M.S.I. Meeting, August 2002, best student oral presentation life sciences section.
- Poster at Institute of Physics Spring Break Conference, April 03, best student poster presentation finalist.
- 2 poster presentations to SPEC 2004 – Shedding New Light on Disease Conference, New York, Jun 04, travel funded by Cancer Research Ireland's Oncology Travel Award 2004.
- Oral presentation to World Conference for Cancer Organisations – WCCO 2005, Burlington Hotel, Dublin, Nov 04.
- Poster presentation at Joint Scientific Meeting, Association of Physical Scientists in Medicine (APSM) and Irish Radiotherapy Physics Group (IRPG), Galway, Feb 2005.
- Oral Presentation to Irish Association for Cancer Research (IACR) and Irish Society for Medical Oncology (ISMO) Joint Conference, Kilkenny, March 05.
- Poster at Institute of Physics Spring Break Conference, March 05, best student poster presentation finalist.
- Oral Presentation to OPTO Ireland, Spectroscopy session, RDS April 05, best oral presentation finalist.
- Invited speaker 1st European DASIM Meeting, Daresbury Synchrotron, July 05.
- Oral Presentation, 28th Microscopical Society of Ireland, Focas Institute, Sep 05, best student oral presentation life sciences.